Essays in Behavioural Economics

by

Shi Zhuo

Thesis Submitted

for the Degree of Doctor of Philosophy in Economics

to the University of Warwick, Department of Economics

September 2021
Contents

List of Figures ......................................................... i
List of Tables ........................................................ ii
Acknowledgements ...................................................... iv
Declaration ............................................................... v
Abstract .................................................................. vi

1 Norms of Apology and Group Cooperation ......................... 1
  1.1 Introduction ...................................................... 1
  1.2 Related Literature ............................................... 3
  1.3 Experiment Design ............................................. 6
  1.4 Results .......................................................... 8
    1.4.1 Treatment Effects ........................................ 9
    1.4.2 The Usage of Apologies ................................ 12
    1.4.3 The Effects of Apologies on Beliefs ....................... 16
    1.4.4 Apology Norms and Cooperative Norms in Groups ...... 19
    1.4.5 The Effects of “Apologies” in the Control ............... 23
  1.5 Discussion ..................................................... 25

2 Ingroup Bias with Multiple Identities: The Case of Religion and Attitudes towards Government Size ...................... 27
  2.1 Introduction ..................................................... 27
  2.2 Methods ........................................................ 29
    2.2.1 Identity Inducement ....................................... 29
    2.2.2 Third-Party Allocation Task .............................. 30
    2.2.3 Post-experiment Questionnaire ........................... 31
    2.2.4 Data Collection ............................................ 32
  2.3 Results ........................................................ 33
    2.3.1 Treatment Effects ........................................ 34
    2.3.2 Possible Explanations .................................... 35
2.4 Discussion ................................................. 40

3 Mindfulness Reduces Information Avoidance ............................. 42
  3.1 Introduction .............................................. 42
  3.2 Literature .................................................. 43
  3.3 Experimental Design ....................................... 44
    3.3.1 Sample Recruitment .................................. 44
    3.3.2 Interventions ........................................ 45
    3.3.3 Procedure ............................................ 46
    3.3.4 Outcomes ............................................. 47
    3.3.5 Regression Specification ............................ 47
  3.4 Results .................................................... 48
    3.4.1 Sample Characteristics .............................. 48
    3.4.2 Effects on Perceived Stress ......................... 48
    3.4.3 Mindfulness and Information Avoidance ........... 48
    3.4.4 Emotion Regulation as a Potential Mechanism .... 50
  3.5 Discussion ................................................ 50

Appendix A (for Chapter 1) ......................................... 52
  A.1 Appendix Figures and Tables ................................ 52
  A.2 Experiment Script ......................................... 59
    A.2.1 Instructions (General Guidelines) .................... 59
    A.2.2 Instructions (Control) ................................ 60
    A.2.3 Screenshots (Practice Round) ....................... 62
    A.2.4 Screenshots (Control) ................................ 62
    A.2.5 Instructions (Public Treatment) ..................... 64
    A.2.6 Screenshots (Public Treatment) ..................... 66
    A.2.7 Instructions (Private Treatment) .................... 68
    A.2.8 Screenshots (Private Treatment) .................... 70
    A.2.9 Instructions (Type Elicitation) ..................... 72
    A.2.10 Screenshots (Type Elicitation) ..................... 74
    A.2.11 Post-Experiment Survey Questions .................. 75

Appendix B (for Chapter 2) ......................................... 78
  B.1 Appendix Figures and Tables ................................ 78
  B.2 Experiment Script ......................................... 86

Appendix C (for Chapter 3) ......................................... 111
  C.1 Information Preference Scale ............................. 111
  C.2 Five Facet Mindfulness Questionnaire ..................... 113
C.3 Perceived Stress Scale ............................................. 114
C.4 Appendix Figures and Tables ................................... 115

References ................................................................. 121
List of Figures

1.1 Determinants of Cooperation over Time ........................................... 9
1.2 Determinants of Apologies ................................................................. 10
1.3 Apologies each round ................................................................. 13
1.4 Group Apology Usage and Group Contributions ............................... 14
1.5 Group Norm of Sincerity and Contributions over Time ....................... 19
1.6 Group Norm of Sincerity and Contributions over Time: Public vs Private ................................................................. 20
1.7 Group Norm of Sincerity and Contributions over Time: Public vs Private ................................................................. 21
2.1 Effects of multiple salient dimensions on ingroup bias by identity dimension ................................................................. 34
2.2 Proportion of participants choosing an equal split (50-50 allocations) by treatment ................................................................. 39
3.1 Intervention Effects on Perceived Stress ............................................. 49
A.1.1 Baseline traits in each treatment .................................................... 52
C.4.1 Number of Participants at Each Session ......................................... 116
C.4.2 Average Session Evaluation ............................................................ 118
C.4.3 Compliance Measure (Average Browser Tab Switches per Session) .... 119
C.4.4 Compliance Measure (Average Difference in Submission Time per Session) ................................................................. 119
C.4.5 Intervention Effects on the Five Facets of Mindfulness Scale ............. 120
## List of Tables

1.1 Treatment Descriptions .................................................. 7
1.2 ...................................................................................... 9
1.3 Treatment Effects on Determinants of Individual Cooperation .... 11
1.4 Beliefs, Targeted Rank and Contributions ............................... 12
1.5 Determinants of Apologies, Further Analysis .......................... 14
1.6 Sending Apologies, Public vs Private: Contributions ............... 16
1.7 Effects of Apologies on Beliefs in Private and Public Treatments .. 18
1.8 Effects of Partner’s Usage of Apologies on Beliefs .................. 20
1.9 Effects of Group Usage of Apologies .................................. 22
1.10 Sending “Apologies” in Control: Contributions ..................... 23
1.11 Effects of “Apologies” in Control: Beliefs ............................ 24
1.12 Effects of Partner’s Usage of “Apologies” on Beliefs in Control ... 24
1.13 Effects of Group Usage of “Apologies” in Control .................. 25

2.1 Breakdown of participants who agreed/disagreed with the religion and government statements .............................................. 33
2.2 Marginal effects of two salient dimensions, relative to one salient dimension, on ingroup bias in each identity dimension ............... 35
2.3 Effects of multiple salient dimensions by identity dimension on survey-measured feelings of belonging .................................. 36
2.4 Marginal effects of two salient dimensions relative to one salient dimension, and ingroup correlation beliefs, on ingroup bias in each identity dimension .................................................. 37
2.5 Marginal effects of feelings of belonging to ingroup, on ingroup bias in each identity dimension ............................................. 38

3.1 Effect of the Treatment on Information Preferences .................. 49
3.2 Effect of the Treatment on Non-Reacting ............................... 50

A.1.1 Apology Treatments and Persistent Cooperative Norms ............ 52
A.1.2 Group Sincerity and Contribution Differences over time .......... 53
A.1.3 Effects of Apologies where Contributions Decrease after Transgression ................................................................. 54
A.1.4 Effects of Apologies where Contributions Decrease after Transgression in Control ........................................ 54
A.1.5 Sending “Apologies” in Control: Contributions, Robustness .... 55
A.1.6 Effects of Apologies in Control: Beliefs, Robustness ............. 55
A.1.7 Effects of Partner’s Usage of Apologies on Beliefs in Control, Robustness ......................................................... 56
A.1.8 Effects of Group Usage of Apologies in Control, Robustness .... 56
A.1.9 Personality Traits and Tendency to Apologise ....................... 57
A.1.10 Personality Traits and Tendency to Sincerely Apologise ........ 57
A.1.11 Importance of Apology over Treatments .......................... 58

B.1.1 Correlation between ingroup bias and survey-measured feelings of belonging ...................................................... 78
B.1.2 Effects of multiple salient dimensions on ingroup correlation beliefs ................................................................. 78
B.1.3 Sample Characteristics ................................................. 79
B.1.4 Marginal effects of two salient dimensions, relative to one salient dimension, on raw ingroup bias in each identity dimension ........ 83
B.1.5 Marginal effects of two salient dimensions relative to one salient dimension, and ingroup correlation beliefs, on raw ingroup bias in each identity dimension ................................. 83
B.1.6 Marginal effects of feelings of belonging to ingroup, on raw ingroup bias in each identity dimension ....................... 83
B.1.7 Marginal effects of two salient dimensions, relative to one salient dimension, on ingroup bias in each identity dimension .......... 84
B.1.8 Effects of multiple salient dimensions by identity dimension on survey-measured feelings of belonging ....................... 84
B.1.9 Marginal effects of two salient dimensions relative to one salient dimension, and ingroup correlation beliefs, on ingroup bias in each identity dimension .................................................. 85
B.1.10 Marginal effects of feelings of belonging to ingroup, on ingroup bias in each identity dimension .............................. 85

C.4.1 Baseline Measures ......................................................... 115
C.4.2 Comparison of Baseline Characteristics of Non-Attritors in Treatment and Control ........................................ 117
Acknowledgements

I would like to thank my supervisor Prof. Daniel Sgroi for his continuous support and guidance throughout my PhD. I would also like to thank other professors, academic and administration staff who have helped me during my years in Warwick. I am fortunate to have received the Leverhulme Bridges Doctoral Scholarship and have benefited a lot from the Director Prof. Thomas Hills and other members of the program, for which I am grateful. I have also had plenty of fruitful and enjoyable interactions with the other PhD students, especially my fellow course mates, to whom I would like to express my gratitude. Finally, my family and friends have given me so much love and support over the years - I am greatly appreciative to them.
Declaration

This thesis is submitted to the University of Warwick in accordance with the requirements of the degree of Doctor of Philosophy in Economics. I declare that it has not been published before or submitted for a degree at another university. Chapter 1 is co-authored with Jonathan Yeo (Nanyang Technological University). Chapter 2 is co-authored with Daniel Sgroi (University of Warwick) and Jonathan Yeo (Nanyang Technological University). Chapter 3 is joint work with Elliott Ash (ETH Zurich), Daniel Sgroi (University of Warwick) and Anthony Tuckwell (University of Warwick). I was involved in all stages of the collaborative projects.

September 2021
Abstract

This thesis explores topics in behaviour economics summarized below.

Chapter 1 investigates norms of apology in the context of group cooperation. Apologies are widely used across various cultures/contexts and have likely evolved to support mutual cooperation in groups. We conduct an experiment examining how apologies affect cooperation in a repeated public goods game. In two separate treatments, participants are given the option to say “I am sorry” either publicly, or privately to group members. In the control, no such option is available. We find that the opportunity to apologise leads to an increase in contributions of 0.37 and 0.61 standard deviations in the private and public treatments respectively. Generally, “norms of apology” exist — participants apologise when contributing less than others, subsequently raising their contributions. Recipients of apologies also believe that apologisers care more about others and will contribute more. In groups, adhering to apology norms is associated with greater salience of cooperative norms — especially in the public treatment where there is common knowledge.

Chapter 2 looks at ingroup bias when multiple dimensions of identities are salient. Group identity is known to exert a powerful socio-psychological influence on behaviour but to date has been largely explored as a uni-dimensional phenomenon. Carefully selecting two politically charged identity dimensions documented to have similar strength and to be largely orthogonal (religious belief and views about government size), we find that priming individuals to consider both dimensions rather than one has a noticeable effect on behaviour. Moving from one to two dimensions can produce a significant increase in ingroup allocations at the expense of fairness to outgroup individuals, although the effect varies as we switch from primarily considering religion to government size. Evidence suggests that the heterogeneity of such effects is related to the degree of “harmony” between groups in the dimensions concerned.

Chapter 3 studies how mindfulness meditation affects information avoidance. Mindfulness meditation has been found to influence various important outcomes such as health, stress, depression, productivity, and altruism. We conduct a randomised-controlled trial to explore a previously untested effect of mindfulness: avoiding information that may cause worry or regret. We find that a relatively short mindfulness treatment (two weeks, 15 minutes a day) is able to induce a statistically significant reduction in information avoidance. Possible mechanisms and policy implications are discussed.
1 Norms of Apology and Group Cooperation

with Jonathan Yeo

1.1 Introduction

Then Pharaoh called for Moses and Aaron in haste; and he said, I have sinned against the Lord your God, and against you. Now therefore forgive, I pray thee, my sin only this once, and intreat the Lord your God, that he may take away from me this death only.

Exodus 10:16–17 — King James Version

Apologising — the act of admitting error and expressing regret — has been omnipresent throughout recent human history, as evident from the millennia old biblical passage. Words of apology are prevalent across cultures and are used in a wide variety of contexts from bilateral to intra-group to inter-group settings. The prevalence of apologies suggests that the words, “I am sorry”, while deceptively simple, may serve an instrumental purpose of greater societal importance.

In particular, apologies likely play an important role in conserving and building social capital (Putnam, 2000), which is important for cooperation in groups. To err is to be human — inevitable mistakes and offenses committed when interacting are always a threat to social cohesion and cooperation. Apologies, and the multitude of social norms associated with them, may have emerged through cultural evolution (Boyd and Richerson, 1985) as a solution to this problem. Relatively, Tavuchis (1991) touts the importance of apologies in being able to paradoxically restore the social order even though they make no amends to the past transgression and are “cheap talk” (Farrell and Rabin, 1996).

How may apologies affect cooperation in groups? We view an apology as more than cheap talk — it involves norms of (i) admitting one’s errors, and (ii) being sincere by making future amends. Since violating these norms of apology risks punishment from others in one’s social group, the norms hence play a role in influencing cooperation.

1The Merriam Webster dictionary defines an apology as an admission of error or discourtesy accompanied by an expression of regret. Almost all the world’s 50 most common languages have a phrase for it. The top 50 most common languages are based on the estimated number of native speakers in the 2007 version of the Swedish encyclopedia Nationalencyklopedin, and comprise $\approx 75\%$ of the world’s population; for more details, see https://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers and https://termcoord.eu/2018/01/im-sorry-in-different-languages/ (accessed March 2020).

2Here, we consider social norms as defined in Crawford and Ostrom (1995): shared understandings about actions that are obligatory, permitted or forbidden.
There are several processes by which they possibly work to reinforce cooperation in the face of inevitable mistakes.

Firstly, by committing to make future amends, apologies lead to more positive beliefs and opinions by receivers of apologies — even in the face of transgressions. Preferences for reciprocity (Rabin, 1993) and conditional cooperation (Fischbacher et al., 2001; Thöni and Volk, 2018) subsequently imply improvements in cooperation. Secondly, having to admit one’s error to others via apologies invokes and amplifies shame (Howell et al., 2012). A desire for social approval (Masclet et al., 2003; Rege and Telle, 2004) means that individuals will likely be more cooperative to avoid making apologies. Thirdly, when apologies are used as intended in groups, they help reaffirm/signal what is considered right or wrong — by making such norms more salient, transgressions can be avoided and cooperation further enhanced.

Furthermore, these mechanisms may interact with whether apologies are made publicly or privately. On one hand, compared to private apologies, public apologies invoke more shame which may discourage their usage; they may also be considered more impersonal (Tavuchis, 1991) even when used — this would reduce their effectiveness. On the other hand, public apologies, by providing public reaffirmation of cooperative norms, could serve as strong reminder to members of the social group to do the “right thing”. It is hence unclear which will be more effective.

In this paper, we conduct an experiment which explores the above issues. We examine how cooperation in groups is affected by the opportunity to make simple interpersonal apologies in a finitely repeated public goods game. Participants are randomly assigned to three different treatments. In two separate apology treatments, participants are given the option to say “I am sorry” either publicly, or privately to group members. In the control treatment, no such option is available. Furthermore, we also elicit full sets of their beliefs about others in order to elucidate the mechanisms by which apologies may improve cooperation.

We find that allowing for apologies significantly improves cooperation, especially when apologies are public information. Compared to the control, contributions are 0.37 and 0.61 standard deviations higher in the private and public treatments respectively. Majority of the cooperation differences can be explained by the positive effects of apologies on beliefs about others. However, consistent with avoiding the need to apologise, participants in the apology treatments also tend to contribute more than what they believe others will contribute.

Conversely, without apologies, higher cooperators become less cooperative as they become more pessimistic about lower cooperators. Relatedly, Ostrom (2000) discusses how solely relying on conditional cooperation can otherwise lead to downward cascades in such events.
Generally, we find that participants exhibit an implicit understanding of what comprise “norms of apology”. When contributing less than others, participants admit their error by apologising. Furthermore, these apologies are sincere and are followed by a rise in contributions. When receiving an apology, recipients are more likely to believe that the sender cares more about others and will contribute more. When these norms of apology are violated, there is evidence of punitive action taken.

Results also show that in groups, norms of sincerity play a larger role than norms of admitting error in raising the salience of cooperative norms. Adhering to norms of sincere apology is strongly associated with greater cooperation in groups, especially in the public treatment where there is common knowledge. Groups with more sincere public apologies exhibit more optimism about others’ contributions and greater cooperation. More sincere private apologies, in contrast, have limited effects.

Placebo tests in the control treatment also show that effects of “synthetic” apologies from individuals in similar positions of transgression do not exhibit any of the above effects. This suggests that the act of apologising itself is important in explaining the improvements in cooperation. Interestingly, responses from the post-experiment questionnaire show that individuals in the control treatment were less likely to want the opportunity to apologies, compared to individuals in the apology treatments. This suggests that without a salient experience, people tend to underestimate the power of apologies. More attention thus need to be paid to apologies and more generally other forms of social etiquette in their role as informal institutions.

The structure of this paper is as follows. Section 2 summarises the main literature related to our paper. Section 3 describes the experimental design and procedures. Section 4 talks about the experimental results, and Section 5 concludes with a discussion.

1.2 Related Literature

There is an extant literature on apologies. In linguistics, psychology and sociology, there has been discussion on the intricacies of apologies and how they play an instrumental role in smoothing relations, for example, see Tavuchis (1991) and Smith
Relatedly, computational models have shown that apologies — via intention recognition and commitment — can lead to strong and lasting cooperation (Han et al., 2013, 2015; Lenaerts et al., 2017). Our study contributes to this body of research by providing causal empirical evidence on how apologies can influence group cooperation.

In economics, there is a small but burgeoning literature on apologies. Several field studies focus on formal apologies from business entities and find mixed effects. More relevant are experimental studies in Fischbacher and Utikal (2013), Ho (2012) and Schniter et al. (2013) which focus on informal apologies and find positive effects. In a static two-player sequential prisoner’s dilemma with punishment, Fischbacher and Utikal (2013) find that the possibility to apologise increases reciprocity. In two player repeated trust games with noise about second mover’s types, Ho (2012) finds that apologies signal types, improving investment. Schniter et al. (2013), in the absence of such noise, find that second movers who send apology-like messages tend to be re-trusted and subsequently return more.

A key contribution of our paper is to show that the effects of informal apologies demonstrated above are important in the dynamics of cooperation in larger groups. To the best of our knowledge, our study is the first to cleanly identify how opportunities to apologise affect cooperation in non-dyadic groups, where effects may be more complex. Furthermore, non-dyadic groups allow us to test whether publicly communicated apologies — in contrast to privately communicated ones — will have stronger effects on cooperation. Which will triumph is not immediately obvious as trade-offs exist between public and private apologies. Relatedly, studies which compare public and private apologies...

---

4 Scher and Darley (1997) identify several key elements of apologies which exhibit independent positive effects on one’s judgements: i) expression of sorrow/regret, ii) taking responsibility for the wrongdoing, iii) providing reparation and iv) committing to behavioural reform. Interestingly, expressing sorrow/regret by saying “I’m sorry” displays the largest effects. Other studies also find positive effects of receiving apologies in both inter-personal and organisational settings. For example, Tabak et al. (2012) find that participants perceive (programmed) transgressors who apologise as more agreeable while Dirks et al. (2011) find that people perceive CEOs who apologise as more repentant, subsequently trusting the transgressors more. There is also research that examines factors which determine apologies. For example, Chrdileli and Kasser (2018); Dunlop et al. (2015); Howell et al. (2012) find associations of certain dispositional traits like guilt and shame with the tendency to apologise while Leunissen et al. (2012) find that perpetrators — especially those with lower dispositional trust — are more willing to apologise when the victim is more willing to forgive. See also Lewicki et al. (2016) for a review of the literature.

5 Abeler et al. (2010) study Ebay sellers, while Ho and Liu (2011) study medical practitioners and find positive effects of apologies. In contrast, Halperin et al. (2019) who study Uber find that the effects of apologies may be limited. Abeler et al. (2010) find that apologies result in higher withdrawal of negative evaulations compared to monetary compensation. Ho and Liu (2011) find that reduced costs of apolologising due to state apology laws result in faster resolutions and lower settlement payments. Halperin et al. (2019) find that apologies by Uber towards their consumers are less effective than monetary compensation in the form of coupons, although they do work when service was only slightly unsatisfactory.
private communication find that effects depend on the context.\textsuperscript{6}

Our study is also broadly related to a vast literature studying group cooperation in the context of public goods games, see Chaudhuri (2011) and Ledyard (1995) for reviews. A robust finding is that in the absence of any external arrangements, cooperation in repeated public goods games decays over time (Ambrus and Pathak, 2011; Chaudhuri et al., 2017; Fischbacher and Gachter, 2010). A body of research examines mechanisms which could improve on this situation.

One subset examines how decentralised punishment after contributing affects cooperation. Such punishment is shown to be effective (Fehr and Gachter, 2000; Nikiforakis and Normann, 2008), although there can be negative anti-social side effects (Dickinson and Masclet, 2015; Herrmann et al., 2008; Nikiforakis, 2008). Another subset examines how communication before contributing affects cooperation. Effects vary depending on the form of communication. While \textit{open ended communication} is effective in improving cooperation (Isaac and Walker, 1988), there are mixed effects of \textit{simple communication} in the form of only a numerical contribution signal (Bochet et al., 2006; Wilson and Sell, 1997).

Our study instead examines how communication after contributing — in the form of apologies — may have its advantages in building group cooperation as it invokes less negative, and more positive feelings.\textsuperscript{7} In our experiment, the opportunity to say “I am sorry”, while verbally simple, also involves possibly complex apology norms. We hence take a middle stance on communication, in contrast to numerical signals which are too simple and open ended messages which are too complex. Particularly relevant, is Bochet and Putterman (2009) who also examine how social norms underlying communication, in the form of promises, can be effective in improving cooperation.\textsuperscript{8} Our social norm approach to apologies broadly relates to the study of informal institutions (North, 1995), relational contracts (Goetz and Scott, 1981) in society and their role in cooperation.

\textsuperscript{6}Förster and Weele (2018) find that public communication can backfire when used to justify less charity and avoid being seen as hypocritical. Chaudhuri and Paichayontvijit (2006) and Chaudhuri et al. (2009), in contrast, find that when advice is made public and is common knowledge, there can be positive effects on coordination.

\textsuperscript{7}To the extent that apologising involves psychological and reputational costs, it can be considered to be a kind of self-inflicted punishment. To the best of our knowledge, there has been no research thus far which examines this kind of self-inflicted punishment.

\textsuperscript{8}Bochet and Putterman (2009) find that in a repeated public goods game, numerical contribution signals combined with non-binding promises can be effective in improving cooperation, at least when punishment is also available.
1.3 Experiment Design

In the experiment, subjects play a standard repeated linear public goods game in groups of four. Groups are randomly assigned to one of three treatments which vary the ability of subjects to send a message of apology at the end of each round.

The Public Goods Game

The primary tool used is a standard linear public goods game which is played in a group of four. Each subject is endowed with 20 coins (points) and has to decide how much to contribute to the public good \( C_i \). In particular, the payoff function is:

\[
\pi_i = 20 - C_i + 0.5 \times \sum_{j=1}^{4} C_j
\]

The private marginal return of contributing (0.5) is less than the cost of contributing (1) and also the social marginal return (2) — rational economic agents should then always contribute nothing (in finite games), which is socially inefficient.

Subjects play a repeated public goods game with the above payoff function for 16 rounds under partner matching. They are paid based on 4 randomly chosen rounds to avoid income effects. Subjects are assigned fixed identities in the group which allows them to track the contributions of each group member over time.

In each round, subjects are asked how much they wish to contribute to a group project (the public good). Furthermore, we also elicit full sets of subjects’ beliefs about other group members. In particular, they are asked about their (i) contribution beliefs — how much each group member will contribute (0 to 20) and (ii) niceness beliefs — how much each group member cares about others (1 to 7). Subsequently, subjects are informed about the contributions of each group member and their own earnings before the next round begins.

Treatments

The three treatments vary the kind of communication which subjects can make at the end of each round — i.e., after observing each group members’ contributions and before the start of the next round, see Table 1.1.

9 All these values are restricted to integer amounts for simplicity. Contribution beliefs are incentivised as in Fischbacher and Gachter (2010): the closer they are to the true contribution, the more they receive. More specifically, if their estimate was exact, they earned 1.5 Coins; if their estimate deviated by one, they earned 1 Coin; if their estimate deviated by two, they earned 0.5 Coins.
In the *control* treatment, which is the baseline, subjects do not have the possibility for such communication. In the *public* and *private* treatments, subjects can instead choose whether to send a message which reads “I am sorry” to their group members at the end of each round.

The public and private treatments differ in terms of whether message(s) can be targeted at specific group member(s). In the *public* treatment, subjects decide whether to send the message to *all group members* — any messages sent are observed by the whole group. By contrast, in the *private* treatment, subjects decide whether to send the message to *each group member* — messages if sent, are observed by only the target recipient.

### Table 1.1: Treatment Descriptions

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Details</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>No communication at end of each round</td>
<td>$n = 76$</td>
</tr>
<tr>
<td>Public</td>
<td>Can send public message: “I am sorry” to whole group after observing contributions</td>
<td>$n = 76$</td>
</tr>
<tr>
<td>Private</td>
<td>Can send private message “I am sorry” to each group member after observing contributions</td>
<td>$n = 80$</td>
</tr>
</tbody>
</table>

**Procedural Details**

The experiment was conducted at the Economics Lab at the University of Warwick between November 2018 and February 2019 and was programmed in oTree (Chen et al., 2016). Participants were recruited using the Warwick SONA system from the student population. In total, there were 13 sessions with 232 participants divided between 58 independent groups. The study was preregistered with the AEA RCT registry (AEARCTR-0003567) and obtained ethical approval for student-led research from the Economics Research Ethics Panel, University of Warwick.

At the beginning of each session, participants were randomly assigned to a group of 4. Each group was then randomly assigned to one of the three treatments. There was hence within session variation in treatments. Participants were told to read through the instructions provided on the web-interface. Subsequently, they had a practice round with simulated group members in order to familiarise them with the user interface. This included a detailed explanation of how payoffs were calculated given the set of contributions in the practice round. They then completed the 16 rounds of the public
At the end of the experiment, their cooperative types were elicited using an incentivised strategy-method procedure as in Fischbacher and Gächter (2010). They were then asked to complete an online post-experiment questionnaire covering demographics, risk preferences, social preferences and their behaviour/experience during the experiment. See Appendix A.2.11 for more details on the survey.

Subjects’ coins were converted into Sterling Pounds at the rate of 100 Coins to £4. Subjects also received a £5 show up/completion fee. The experiment lasted about 75 minutes and subjects earned an average of £11.

**Outcome Variables**

The experimental setup allows for the measurement of several outcome variables of interest. Our main variable of interest is individual contribution, which measures cooperation. In line with our earlier discussion, we further focus on several possible determinants of cooperation — 1) contribution beliefs, since fairness preferences (Fehr and Schmidt, 1999) imply a desire to reciprocate others contributions; 2) niceness beliefs since reciprocal kindness (Rabin, 1993) implies a desire to reciprocate nice/kind individuals; 3) targeted contribution rank, since social considerations imply a desire not to rank too low.

1.4 Results

In this section, we will begin with an examination of the aggregate treatment effects of public and private apologies. This will be followed by a closer look at the individual usage of apologies and how it affects the evolution of cooperation from round to round. Finally, we examine how group usage of apologies influences cooperation.

Our primary tool will be regression analyses at various levels. At the individual level, we have 232 observations; at the individual-round level, we have 232 × 16 = 3712 observations; at the individual-partner-round level, we have 232 × 16 × 3 = 11136

10 Gender was approximately balanced across the treatments. There were however small differences in age — participants in the private treatment were slightly younger. See Figure A.1.1. Most participants also displayed a good understanding of the experiment — 98.7% indicated that they understood the main experiment. Controlling for age or understanding in all analyses does not change the results. All participants were provided with paper and a pen to record anything down if they wished. For more details of the experimental design, see Appendix A.2.

11 We focus on these 3 factors since being inter-personal factors, they are likely to be influenced by apologies, in contrast to personal factors like altruism. Measuring participants’ beliefs about each group member allows us to calculate their targeted rank in each round. More precisely, targeted rank is the number of others for which personal contributions < contribution beliefs. A lower number for this is referred to as a higher rank.
observations. In all regressions, we include session fixed effects. For regressions at the panel level, we also include round fixed effects. Since the statistically independent unit is the group, we always cluster standard errors at the group level.

1.4.1 Treatment Effects

Individual Cooperation

Figure 1.1 plots participants’ contributions in each treatment over time. It shows an improvement in cooperation with the opportunity to apologise, especially in the public treatment. Interestingly, contributions are already significantly higher in the apology treatments during the first round, \( p = 0.012 \). Observe that a gap grows between the apology treatments and the control, with a larger gap emerging in the public treatment. Figure 1.1 also shows that while there are end-round effects, effects of apologies still persist to some extent in the final round.\(^\text{12}\)

![Figure 1.1](image)

<table>
<thead>
<tr>
<th>Table 1.2</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public</td>
<td>4.232***</td>
</tr>
<tr>
<td></td>
<td>(1.005)</td>
</tr>
<tr>
<td>Private</td>
<td>2.551**</td>
</tr>
<tr>
<td></td>
<td>(1.113)</td>
</tr>
<tr>
<td>Constant</td>
<td>8.937***</td>
</tr>
<tr>
<td></td>
<td>(1.352)</td>
</tr>
</tbody>
</table>

Session, Round ✓

Fixed Effects

Observations 3712

\(* 0.10 \hspace{0.5em} ** 0.05 \hspace{0.5em} *** 0.01. OLS with individual-round level observations. Standard errors in parentheses clustered at group level.\)

The regression in Table 1.2 confirms that apologies improve cooperation. Compared to the control, contributions are 4.23 points (0.61 standard deviations) higher in the public treatment, \( p < 0.001 \), and 2.55 points (0.37 standard deviations) higher in the private treatment, \( p = 0.026.\(^\text{13}\) In addition, common knowledge looks to be important: contributions are significantly higher in the public treatment compared to the private

\(^{12}\)Results from the post-experiment elicitation further indicate that cooperative types are more prevalent in the apology treatments. Table A.1.1 shows that last round contributions and unconditional cooperation are higher in the apology treatments, especially in the public treatment. There are also differences between the control and apology treatments in conditional cooperation types. In the public treatment, conditional cooperation has a higher base level, while in the private treatment, conditional cooperation is slightly “steeper”. This suggests that apologetic communication does raise cooperative tendencies to some extent.

\(^{13}\)The positive effect of apologies on cooperation does not come at the expense of a rise in inequality. In fact, the standard deviation of contributions in each round are lower on average under apologies, albeit insignificantly so, \( p = 0.294.\)
treatment, $p = 0.076$. As cooperation is socially efficient, effects on social welfare also follow similar trends — compared to the control, average payoffs in groups are 0.81 (0.49) standard deviations higher in the public (private) treatment.

Several comparisons hint at why this may be the case. Transgressors — who contribute strictly less than some other — tend to be more “repentant” in the apology treatments. 46% of transgressors raise their contributions in the control compared to 59% and 54% in the public and private treatments respectively. Non-transgressors — who contribute weakly more than all others — also tend to be more “forgiving” in the apology treatments. 45% of non-transgressors lower their contributions in the control compared to 27% and 39% in the public and private treatments respectively.\textsuperscript{14} We will come back to how apologies lead to this in subsequent sections.

**Beliefs and Targeted Ranks**

Figure 1.2 plots beliefs and targeted contribution ranks in each treatment over time. In the first round, there are no significant differences in contribution or niceness beliefs between the apology treatments and the control, $p = 0.779$ and $p = 0.352$ respectively. However, targeted ranks are significantly higher in the first round, $p = 0.003$, which is consistent with the higher first round contributions despite no differences in contribution beliefs. Generally, the trends of contribution beliefs, niceness beliefs, and targeted contribution ranks broadly resemble that of actual contributions in Figure 1.1; albeit with less difference in the public and private treatments for targeted ranks.

Figure 1.2: Determinants of Cooperation over Time

\textsuperscript{14}For transgressors, t-tests of differences with the control, $p = 0.013$ and $p = 0.137$ for the public and private treatments respectively. For non-transgressors, t-tests of mean differences with the control, $p = 0.042$ and $p = 0.466$ for the public and private treatments respectively.
Regressions in Table 1.3 confirm that contribution and niceness beliefs are higher with apologies, especially public ones. It shows that relative to the control, contribution beliefs are 4.46 points (0.78 standard deviations) higher under public apologies, \( p < 0.001 \), compared to 2.30 points (0.40 standard deviations) higher under private apologies, \( p = 0.046 \). Niceness beliefs are 0.99 points (0.60 standard deviations) higher under public apologies, \( p < 0.001 \), compared to 0.54 points (0.33 standard deviations) higher under private apologies, \( p = 0.051 \). Beliefs are significantly higher in the public treatment compared to the private treatment, \( p = 0.026 \) and \( p = 0.054 \) for contribution beliefs and niceness beliefs respectively.

Table 1.3: Treatment Effects on Determinants of Individual Cooperation

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Contribution Beliefs</th>
<th>Niceness Beliefs</th>
<th>Targeted Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public</td>
<td>4.463*** (0.996)</td>
<td>0.989*** (0.232)</td>
<td>-0.188 (0.120)</td>
</tr>
<tr>
<td>Private</td>
<td>2.297** (1.128)</td>
<td>0.540* (0.271)</td>
<td>-0.180* (0.104)</td>
</tr>
<tr>
<td>Constant</td>
<td>9.169*** (1.233)</td>
<td>4.510*** (0.230)</td>
<td>2.156*** (0.184)</td>
</tr>
<tr>
<td>Session, Round Fixed Effects</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Observations</td>
<td>3712</td>
<td>3712</td>
<td>3712</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. OLS regressions with individual-round level observations. Standard errors in parentheses clustered at the group level. Contribution and Niceness beliefs are averaged over all partners. Contribution beliefs are on a scale from 0 to 20 while niceness beliefs are on a scale from 0 to 7.

Table 1.3 also shows that with apologies, average targeted ranks are approximately 0.18 – 0.19 higher than in the control. Pooling the apology treatments, the effect is significant, \( p = 0.074 \); there are however no big differences between public and private apologies, \( p = 0.937 \). There is thus evidence that participants try to avoid the need to apologise should one’s contribution rank be too low.\(^{15}\)

Mediating Effects of Beliefs and Targeted Ranks

The above analysis shows strong concordance of the treatment effects on beliefs and targeted rank with that on actual contributions. This suggests that beliefs and targeted rank indeed play a mediating role in influencing cooperation. Examining this, Table 1.4, Columns 2 to 4 show that these variables are all significantly correlated with higher contributions and reduce the magnitude of differences between treatments. Table 1.4,\(^{15}\)

\(^{15}\)This may have arisen spuriously because increases in contributions from apologisers to amend for their mistakes mechanically lead to higher targeted ranks on aggregate in the apology treatments. However, having apologised in the past is associated with a 0.272 lower targeted rank on average, which rules out this explanation \( (p = 0.005) \). Furthermore, as mentioned, in the first round, targeted rank is already higher with apologies: contributions are significantly higher \( (p = 0.012) \) even though contribution beliefs are not significantly different \( (p = 0.779) \).
Column 5 shows that including all of these variables can explain almost all of the differences between treatments.

Table 1.4: Beliefs, Targeted Rank and Contributions

<table>
<thead>
<tr>
<th></th>
<th>Dep Var: Contribution</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public</td>
<td></td>
<td>4.232***</td>
<td>1.059***</td>
<td>2.680***</td>
<td>3.641***</td>
<td>0.156</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.005)</td>
<td>(0.353)</td>
<td>(0.773)</td>
<td>(0.927)</td>
<td>(0.276)</td>
</tr>
<tr>
<td>Private</td>
<td></td>
<td>2.551**</td>
<td>0.918**</td>
<td>1.703**</td>
<td>1.984*</td>
<td>0.165</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.113)</td>
<td>(0.359)</td>
<td>(0.844)</td>
<td>(1.028)</td>
<td>(0.266)</td>
</tr>
<tr>
<td>Contribution Beliefs</td>
<td></td>
<td>0.711***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.042)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Niceness Beliefs</td>
<td></td>
<td></td>
<td>1.569***</td>
<td></td>
<td></td>
<td>0.422**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.159)</td>
<td></td>
<td></td>
<td>(0.198)</td>
</tr>
<tr>
<td>Targeted Rank</td>
<td></td>
<td>3.152***</td>
<td></td>
<td></td>
<td>3.303***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.223)</td>
<td></td>
<td></td>
<td>(0.198)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td>8.937***</td>
<td>2.419***</td>
<td>1.860</td>
<td>3.125***</td>
<td>-5.299***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.352)</td>
<td>(0.697)</td>
<td>(1.350)</td>
<td>(1.027)</td>
<td>(0.797)</td>
</tr>
<tr>
<td>Session, Round Fixed Effects</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td>3712</td>
<td>3712</td>
<td>3712</td>
<td>3712</td>
<td>3712</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. OLS regressions with individual-round level observations. Standard errors in parentheses clustered at the group level. Contribution and Niceness beliefs are averaged over all partners. Contribution beliefs are on a scale from 0 to 20 while niceness beliefs are on a scale from 0 to 7.

A mediation analysis where beliefs and targeted rank mediate the effects of the treatments on contributions indicates that beliefs have the most explanatory power. Contribution beliefs and niceness beliefs together explain 81% (70%) of the differences between the control and public (private) treatment compared to 15% (23%) for targeted rank. Almost all — 99% — of the differences between the private and public treatment can be explained by beliefs alone. There is thus strong evidence that almost all differences in contributions between treatments can be explained by how public apologies and private apologies affect beliefs and targeted ranks.

In what follows, we first document how apologies are used and what comprises the norm of apologies. Subsequently, we examine how the sending and receiving of apologies affects these beliefs given their importance in determining contributions as seen from above.

1.4.2 The Usage of Apologies

In the experiment, apologies are widely used: a majority of subjects (92%) sent an apology at least once. Figure 1.3 displays the number of apologies sent in the public and private treatments over time. It shows that apologies sent are initially quite similar across the treatments. However, a gap emerges over time as apology rates fall in the public treatment while remaining roughly constant in the private treatment. On
average, the mean number of apologies sent is 0.73 in the public treatment and 0.92 in the private treatment. This likely reflects the higher levels of coordination which emerge in the public treatment.16

In what follows, we show that there is much consistency in how apologies are used — this suggests that the act of saying “I am sorry” involves social norms of apology.

**When are apologies used?**

One aspect of apologies discussed earlier was that it involved a norm of admitting one’s error when a transgression occurs. Results are consistent with participants having a common understanding of what constitutes a transgression. Figure 1.4, Panel A illustrates that individuals tend to send apologies when contributing less than another. Figure 1.4, Panel B shows that on average, individuals with lower contribution rank tend to apologise more.

Table 1.5 provides a closer examination of how these factors influence the choice of whether to apologise. It shows that individuals with lower contribution ranks indeed tend to apologise more on average. Furthermore, when not the highest rank, contributing less relative to the other raises the likelihood of apologising — more so at higher ranks. This is especially so in the public treatment where the contribution difference only matters when one’s rank is the second highest. The lower relevance of contribution differences in the public treatment likely reflects the inability to target apologies compared to the private treatment.

16 In a t-test of differences in mean apologies sent, \( p = 0.128 \). An OLS regression of apology rates on a private treatment dummy, round and their interaction indicates a slightly less negative round effect in the private treatment (\( p = 0.152 \)). One might wonder why there are still apologies in later rounds. We surmise that unlike games where players have fixed types and signalling is only needed once with no noise, temptation and error can lead to (accidental) violation of norms which require apologies to maintain constant coordination. This may be especially relevant when full coordination has not been attained.
Figure 1.4: Determinants of Apologies

Data for this involves individual-partner-round level observations. Error bars are 95% confidence intervals with standard errors clustered at the group level. For the public treatment, an apology is considered to be sent to all members of the group. The apology rate is the proportion of partners for which an apology is sent. Panel A splits observations by whether one has contributed strictly less, or weakly more than a partner. Panel B splits observations by one’s contribution rank — the number of members who contributed strictly more than me + 1.

Table 1.5: Determinants of Apologies, Further Analysis

<table>
<thead>
<tr>
<th>Dep Var: Made Apology</th>
<th>Public</th>
<th>Private</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contribution Difference</td>
<td>0.021</td>
<td>0.041*</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Rank=2</td>
<td>0.975*</td>
<td>0.968***</td>
</tr>
<tr>
<td></td>
<td>(0.546)</td>
<td>(0.250)</td>
</tr>
<tr>
<td>Rank=3</td>
<td>2.134***</td>
<td>1.355***</td>
</tr>
<tr>
<td></td>
<td>(0.550)</td>
<td>(0.337)</td>
</tr>
<tr>
<td>Rank=4</td>
<td>2.909***</td>
<td>1.880***</td>
</tr>
<tr>
<td></td>
<td>(0.575)</td>
<td>(0.474)</td>
</tr>
<tr>
<td>Rank=2 × Contribution Difference</td>
<td>-0.073*</td>
<td>-0.240***</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>Rank=3 × Contribution Difference</td>
<td>-0.048</td>
<td>-0.136***</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td>(0.031)</td>
</tr>
<tr>
<td>Rank=4 × Contribution Difference</td>
<td>-0.027</td>
<td>-0.072</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.049)</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.826***</td>
<td>-2.128***</td>
</tr>
<tr>
<td></td>
<td>(0.568)</td>
<td>(0.487)</td>
</tr>
</tbody>
</table>

Session, Round Fixed Effects ✓ ✓

Observations 3648 3840

* 0.10 ** 0.05 *** 0.01. Logit regressions with individual-partner-round level observations. Standard errors in parentheses clustered at the group level. Contribution Difference = Own Contribution - Other’s Contribution.
While the above shows that apologies are more likely to be used when contributing less than others, Figure 1.4, Panel B also shows that a non-zero fraction of top contributing individuals apologise. Survey responses indicate that these individuals use apologies differently — they were less likely to state contributing less than others as the sole reason for apologising, but had other reasons like incorrect estimations of others’ contributions, guilt tripping others, or trying to be encouraging. As these do not conform to standard reasons for apologising — we refer to them as non-standard apologies hereinafter. We abstract from these non-standard apologies and focus on the effects of apologies made when not at the highest contribution rank in our analysis.17

Are apologies cheap talk or are they sincere?

A second aspect of apologies in our discussion was that it involved a norm of being sincere and making amends. For participants at the non-highest rank, 44% – 51% who do not apologise raise their contributions, compared to 59% – 72% of those who apologise. This suggests that on average, apologisers indeed live up to their apologies. A more detailed analysis of how sending apologies affects contributions ($C_{it}$) is performed by estimating the following econometric model.18

$$C_{it} = \beta_0 + \beta_1 \text{Apology}_{it(t-1)} + \beta_2 C_{i(t-1)} + \beta_3 C_{-i(t-1)} + \beta_4 Z_{it(t-1)} + W_{it} + e_{it}$$

Table 1.6 shows that apologisers tend to raise their contributions in the next round, however there is no difference between public and private apologies. Aggregating the two apology treatments, Column 1 shows that controlling for past contributions relative to the group, each apology sent is associated with a 0.69 point increase in contributions on average.

---

17Amongst those who had ever made a non-standard apology, 24% specifically indicated contributing less than others as the sole reason for apologising compared to 65% amongst those who only made standard apologies. The difference is significant, $p < 0.001$. 47 out of 296 apologies (16%) in the public treatment and 260 out of 1182 apologies (22%) in the private treatment are non-standard apologies — these are not treated as actual apologies in our analysis. While we exclude non-standard apologies for brevity in our subsequent main analyses, including them as controls in regressions does not qualitatively change our results.

18Type I tobit regressions are used since contributions are censored from above and below. $C_{i(t-1)}$, $C_{-i(t-1)}$ control for own and others’ past contributions since relative contributions are likely to affect adjustments. $Z_{it(t-1)}$ is a set of past round controls including one’s contribution rank and a dummy for coordination (full contributions in the group). $W_{it}$ includes session and round fixed effects.
Comparing the public and private apology treatments, Column 2 shows that there does not seem to be a statistically significant difference in effects of sending apologies. There is however an overall tendency to be more positive in decisions to contribute for the public treatment (see Private coefficient) — this will be examined later in Section 1.4.4.

### 1.4.3 The Effects of Apologies on Beliefs

Here, we analyse the effects of apologies on beliefs using the data set consisting of individual PARTNER-round level observations — subject i’s belief (B) about each group member j in round t. In particular we estimate variations of the following econometric model:

\[
B_{ijt} = \alpha_0 + \alpha_1 \text{Apology}_{ij(t-1)} + \alpha_2 B_{ij(t-1)} + \alpha_3 C_{jt(t-1)} + \alpha_4 C_{jt(t-1)} + \alpha_5 Z_{ij(t-1)} + W_{it} + u_{ijt}
\]

---

19 Results are also similar if using dummies for the number of apologies sent. Aggregating the apology treatments, sending 1, 2 or 3 apologies is associated with higher contributions, \( p = 0.093, p = 0.001, p = 0.001 \) respectively. The effects of sending 3 private apologies is smaller than that of sending 3 public apologies, but this difference is not significant, \( p = 0.716 \).

20 Type I tobit regressions are used since beliefs are censored from above and below. \( C_{jt(t-1)} \) control for past contribution of the partner and others besides the partner. \( Z_{ij(t-1)} \) is a set of past round controls including one’s contribution rank, the partner’s contribution rank and a dummy for coordination (full group contributions). \( W_{it} \) includes session and round fixed effects.
How does receiving apologies affect one’s beliefs?

Previously, we found that apologies were not cheap talk. Results indicate that recipients of apologies do not treat them as cheap talk either. Table 1.7 shows that recipients of apologies believe that those who apologise will contribute more and feel that they care more about others. Columns 1 and 3 show that aggregating both treatments, receiving an apology is on average associated with a 1.15 point rise in contribution beliefs, \( p = 0.002 \), and a 0.22 point rise in niceness beliefs, \( p < 0.001 \).

Since these received apologies are common knowledge in the public treatment, one might expect its effect on contribution beliefs to be greater due to higher order effects through third parties.\(^{21}\) Column 2 shows that there do not seem to be significant differences between the private and public treatment, \( p = 0.901 \). One reason could be that these higher order effects are too weak to detect. Column 4 also shows no significant differences for niceness beliefs, \( p = 0.914 \). Results here are consistent with Table 1.6 where similarly, no differences in effects of sending apologies were observed.

How does sending apologies affect one’s beliefs?

Given that apologies sent are not cheap, if others believed in them as shown above and raised contributions, a corollary would be that sending apologies should also have an indirect effect on one’s contribution beliefs. Table 1.7 shows evidence consistent with this. Aggregating the apology treatments, Column 1 shows that on average, having sent an apology is associated with a 0.40 point rise in contribution beliefs about another.

Since apologies sent are common knowledge in the public treatment, its effect on contribution beliefs might again be greater due to higher order effects through third parties.\(^{22}\) Column 2 shows that like before, there is no significant difference between the public and private treatments in this regard, \( p = 0.565 \). Interestingly, Column 4 shows that sending a private apology, but not a public apology, raises ones niceness beliefs about another, \( p = 0.042 \).

---

\(^{21}\) Receiving an apology in the public treatment implies that third parties receive an apology as well. Hence, third parties’ beliefs about the sender partner’s contribution also rise. This subsequently raises the third parties’ contributions. As a result, the sender partner’s contribution beliefs and thus contributions should also be expected to increase. In contrast, this should not hold for niceness beliefs. Receiving an apology in the private treatment has no such effects.

\(^{22}\) Sending an apology in the public treatment implies that third parties receive the apology as well. Hence, third parties’ beliefs about one’s contribution also rise. This subsequently raises their contribution beliefs and contributions. As a result, the receiver partner’s contribution beliefs and hence contributions should also be expected to increase. In contrast, this should not hold for niceness beliefs. Sending an apology in the private treatment has no such effects.
Are there effects of third-party public apologies?

Previously, we tested whether common knowledge of apologies received and sent could lead to stronger effects on contribution beliefs in the public treatment — due to indirect effects through third parties — finding no significant effects. Here, we examine whether common knowledge of third party apologies received by one’s partner could directly lead to stronger effects on contribution beliefs in the public treatment.\(^{23}\)

Table 1.7, Columns 2 and 4 show some evidence consistent with this, albeit less precisely estimated. Compared to the the private treatment, each apology a partner receives from third parties in the public treatment is associated with a 0.49 point rise in contribution beliefs about the target, \(p = 0.100\). In contrast, there is no sizeable difference for niceness beliefs, \(p = 0.860\).

Table 1.7: Effects of Apologies on Beliefs in Private and Public Treatments

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Contribution Belief</th>
<th>Niceness Belief</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apology Sent</td>
<td>0.402* (0.243)</td>
<td>0.074 (0.095)</td>
</tr>
<tr>
<td>Apology Received</td>
<td>1.145*** (0.369)</td>
<td>0.222*** (0.061)</td>
</tr>
<tr>
<td>No. of 3rd Party Apologies Received by Partner</td>
<td>-0.005 (0.173)</td>
<td>-0.093** (0.037)</td>
</tr>
<tr>
<td>Private</td>
<td>-0.860*** (0.328)</td>
<td>-0.042 (0.091)</td>
</tr>
<tr>
<td>Private × Apology Sent</td>
<td>0.226 (0.392)</td>
<td>0.224*** (0.110)</td>
</tr>
<tr>
<td>Private × Apology Received</td>
<td>0.081 (0.649)</td>
<td>0.014 (0.127)</td>
</tr>
<tr>
<td>Private × No. of 3rd Party Apologies Received by Partner</td>
<td>-0.490 (0.298)</td>
<td>-0.014 (0.077)</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.136 (0.990)</td>
<td>1.549*** (0.214)</td>
</tr>
<tr>
<td>Past Round Controls ✓ ✓ ✓ ✓</td>
<td>✓ ✓ ✓ ✓</td>
<td>✓ ✓ ✓ ✓</td>
</tr>
<tr>
<td>Observations 7020 7020 7020 7020</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\* 0.10 ** 0.05 *** 0.01. Tobit regressions with individual-partner-round level observations. Standard errors in parentheses clustered at the group level. Controls: Past contribution of partner, Past contribution of others besides partner, Past contribution rank, Partner’s past contribution rank, Past coordination, Past belief. Apologies are restricted to those made at non-highest rank.

Noticeably, in the belief regressions, there is also generally greater optimism in the formation of contribution beliefs in the public treatment, compared to the private treatment (see Private coefficient) — this will be examined in Section 1.4.4.

\(^{23}\)Public apologies received from third parties by one’s partner should raise his/her contribution beliefs and consequently contributions. Since such third party apologies are observed in the public treatment, but not in the private treatment, one’s contributions beliefs about the partner should increase in the public treatment. In contrast, apologies from third parties should not affect niceness beliefs about the partner, whether in the public or private treatment.
1.4.4 Apology Norms and Cooperative Norms in Groups

So far, we have focused on how individual apologies affect cooperation. However, apologies may have effects beyond the individual level. How apologies are used in groups may also act as a signal which reaffirms a relational contract to cooperate. The reliability of this signal can be supported by the actuality or threat of social sanctions that can more than offset the short-term benefits of cheating and deception (Schniter et al., 2013). We find several pieces of evidence consistent with this.

Figure 1.5: Group Apology Usage and Group Contributions

The slope coefficients for Panels A, B and C are respectively: \(-0.166 \ (p = 0.314)\), \(2.26 \ (p = 0.543)\) and \(14.6 \ (p < 0.001)\). The group norm of admitting error is measured as the proportion of times in the group where an apology is made in cases where a partner has a strictly higher contribution. The group norm of sincerity is measured as the proportion of times in the group where an apology is followed by a strict increase in contributions. Both standard and non-standard apologies are considered for sincerity since we are interested in the overall observed apology norm.

Firstly, we find that norms of “proper” apology usage within groups are associated with stronger cooperative norms. Figure 1.5 examines two aforementioned aspects of apologies: (i) whether they are made as admission of errors, and (ii) whether they are sincere. It shows that the two aspects are weakly correlated and that apologising sincerely is more strongly associated with stronger cooperative norms — as measured by average contributions across all rounds. In a regression of average group contributions on group norms of sincerity and admitting error, \(\beta_{\text{sincere}} = 15.376 \ (p < 0.001)\), \(\beta_{\text{admit error}} = 4.808 \ (p = 0.090)\). This suggests that an admission of one’s error is not enough; apologies also need to be sincere.

Figure 1.6 provides further evidence that the above correlation of group sincerity and contributions does not just reflect heterogeneity in cooperative types. It shows that while contributions are initially quite similar across the upper and lower tertiles of group sincerity, a gap between them opens up over time.\(^{24}\) This would be consistent

\(^{24}\) Regressions of individual contributions on group sincerity, interacted with round fixed effects, also show that positive effects of group sincerity emerge over time, see Table A.1.2, Column 1.
with proper usage of apologies making the norm of “what is right” more salient, thus reinforcing cooperation.

Secondly, there is evidence of punitive action when norms of apology are not followed. Table 1.8 shows that contribution beliefs are more optimistic when one’s partner uses apologies to admit their errors and these apologies are sincere. This effect is however absent for niceness beliefs, perhaps because these are more spontaneous in nature. Given conditional cooperation, this implies that insincerity will be “punished” with lower future cooperation. Furthermore, we also find evidence that if there is no intention to make amends, apologising can even backfire relative to not apologising. The decline in cooperation towards the end of experiment could indicate a breakdown in the relational contract due to the lower consequences of such sanctions at that stage.

Table 1.8: Effects of Partner’s Usage of Apologies on Beliefs

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Contribution Belief</th>
<th>Niceness Belief</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partner’s Past Sincerity</td>
<td>0.814**</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>(0.326)</td>
<td>(0.075)</td>
</tr>
<tr>
<td>Partner’s Past Admission of Errors</td>
<td>0.737**</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>(0.366)</td>
<td>(0.071)</td>
</tr>
<tr>
<td>Private</td>
<td>-0.618</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>(0.469)</td>
<td>(0.100)</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.145*</td>
<td>1.423***</td>
</tr>
<tr>
<td></td>
<td>(1.259)</td>
<td>(0.313)</td>
</tr>
<tr>
<td>Past Round Controls and Apologies</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Session, Round Fixed Effects</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>4886</td>
<td>4886</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Tobit regressions with individual-partner-round level observations. Subset of observations where transgressions and apologies have been observed in the past. Standard errors in parentheses clustered at the group level. Past Admission of Errors is the proportion of rounds in the past where a partner apologises when contributing strictly less. Past Sincerity is the proportion of rounds in the past where a partner strictly increases contributions after apologising.

Controls: Past contribution of partner, Past contribution of others besides partner, Past contribution rank, Partner’s past contribution rank, Past coordination, Past belief, Standard apologies received and sent (× treatments).

25 For cases where a partner was not a top contributor, and subsequently reduced contributions, Table A.1.3 shows that having made an (insincere) apology in contrast to no apology is associated with a reduction in the contribution beliefs about the partner, albeit imprecisely estimated, \( p = 0.169 \).
Differences between the Public and Private Treatments

Previously, we showed that adherence to norms of apology in the group — especially sincerity — is associated with stronger relational contracts to cooperate. In the public treatment, adherence to such norms is common knowledge, likely easing the formation of a relational contract. This might then explain why cooperative norms are stronger under public apologies compared to private apologies despite minimal differences in their direct effects as shown in Section 4.2 and 4.3. We find that differences between the public and private treatment in responses to group norms of sincerity are consistent with this hypothesis.

Figure 1.7 shows that stronger cooperative norms emerge more quickly in the public treatment when group norms of sincerity are strong. In both the public and private treatments, a gap between the upper and lower tertiles of group sincerity opens up over time. However, in the private treatment, the trend difference is much less steep.  

Figure 1.7: Group Norm of Sincerity and Contributions over Time: Public vs Private

Table 1.9 instead shows regressions analogous to past regressions which examined the evolution of cooperation and beliefs in response to individual apologies (Tables 1.6 and 1.7), additionally controlling for effects of group norms of apology in each treatment. Columns 1 and 2 show that in the public treatment, there is greater optimism in the formation of beliefs when apologies are used sincerely within the group. In the private treatment, by contrast, the effects are weaker. Column 3 also shows similar effects for decisions to contribute. While previous analysis noted an overall optimism in the public treatment (a negative Private coefficient), this effect disappears once differential responses to group sincerity are controlled for.

---

26 Regressions of individual contributions on group sincerity, interacted with round fixed effects for each treatment also show that positive effects of group sincerity emerge earlier in the public treatment, see Table A.1.2, Columns 2 and 3.

27 The differences in the interactions are significant, \( p = 0.087 \), \( p = 0.001 \), \( p = 0.076 \) respectively for Columns 1 to 3.
Table 1.9: Effects of Group Usage of Apologies

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Contribution Beliefs</th>
<th>Niceness Beliefs</th>
<th>Contributions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public × Group Norm of Sincerity</td>
<td>4.455***</td>
<td>0.800***</td>
<td>8.360***</td>
</tr>
<tr>
<td></td>
<td>(1.745)</td>
<td>(0.223)</td>
<td>(2.276)</td>
</tr>
<tr>
<td>Private × Group Norm of Sincerity</td>
<td>0.369</td>
<td>0.022</td>
<td>2.673</td>
</tr>
<tr>
<td></td>
<td>(1.764)</td>
<td>(0.196)</td>
<td>(2.100)</td>
</tr>
<tr>
<td>Public × Group Norm of Admitting Error</td>
<td>1.093</td>
<td>-0.419</td>
<td>-0.643</td>
</tr>
<tr>
<td></td>
<td>(0.802)</td>
<td>(0.271)</td>
<td>(1.119)</td>
</tr>
<tr>
<td>Private × Group Norm of Admitting Error</td>
<td>0.904</td>
<td>-0.052</td>
<td>0.808</td>
</tr>
<tr>
<td></td>
<td>(2.310)</td>
<td>(0.345)</td>
<td>(2.769)</td>
</tr>
<tr>
<td>Private</td>
<td>1.759</td>
<td>0.307</td>
<td>2.378</td>
</tr>
<tr>
<td></td>
<td>(2.229)</td>
<td>(0.236)</td>
<td>(2.971)</td>
</tr>
<tr>
<td>Constant</td>
<td>-4.079***</td>
<td>1.281***</td>
<td>-3.357**</td>
</tr>
<tr>
<td></td>
<td>(1.180)</td>
<td>(0.241)</td>
<td>(1.408)</td>
</tr>
<tr>
<td>Past Round Controls and Apologies</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Session, Round Fixed Effects</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>7020</td>
<td>7020</td>
<td>2340</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Columns 1 - 2: Tobit regressions with individual-partner-round level observations. Column 3: Tobit regression with individual-round level observations. Standard errors in parentheses clustered at the group level. Group norm of admitting error is measured as the proportion of times in the group where an apology is made in cases where a partner has a strictly higher contribution. Group norm of sincerity is measured as the proportion of times in the group where an apology is followed by a strict increase in contributions.

 Controls (Columns 1 - 2): Past contribution of partner, Past contribution of others besides partner, Past contribution rank, Partner’s past contribution rank, Past coordination, Past belief, Standard apologies received and sent. Controls (Column 3): Past contribution, Avg of others’ past contributions, Past contribution rank, Past Coordination, Standard apologies sent.
1.4.5 The Effects of “Apologies” in the Control

The results thus far show that apologies in the public and private treatments are not cheap talk and have positive effects on beliefs and contributions. Furthermore, there is an effect beyond individual apologies when groups follow apology norms. Here, we provide further evidence that these relationships are not spurious by conducting placebo tests using data from the control treatment. In particular, we show that effects of potential choices to apologise in the control in similar situations differ from our previous results and are largely non-existent.\textsuperscript{28}

An issue is that potential choices to apologise are not observed in the control. To deal with this, we simulate “synthetic” apologies in the control based on how participants apologised in the public and private treatment. In particular, logit regressions similar to Table 1.5 are performed for pooled apology choices in the public and private treatments. The coefficients are then used to predict the probabilities of apologies given observed values of the independent variables in the control. These probabilities are then used to simulate apologies in the control.\textsuperscript{29} Regressions similar to that in the previous sections are then conducted for the control.

Table 1.10: Sending “Apologies” in Control: Contributions

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Apologies Sent</td>
<td>-0.189</td>
</tr>
<tr>
<td></td>
<td>[0.458]</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.502*</td>
</tr>
<tr>
<td></td>
<td>[0.075]</td>
</tr>
<tr>
<td>Past Round Controls</td>
<td>✓</td>
</tr>
<tr>
<td>Session, Round Fixed Effects</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>1140</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Tobit regressions with individual-round level observations. Standard errors clustered at the group level. P-values in square parentheses. Estimates are based on the means of 2000 sets of apology simulations in the control group. Controls: Past contribution, Avg of others’ past contributions, Past contribution rank, Past Coordination. Apologies are restricted to those made at non-highest rank.

The placebo tests indicate no significant effects on contributions and beliefs for the effects of individual apologies. Unlike our earlier findings, Table 1.10 shows that sending “synthetic” apologies is not associated with higher contributions while Table 1.11 shows that receiving “synthetic” apologies is not associated with higher beliefs.\textsuperscript{28}

\textsuperscript{28}Since our analysis is mainly correlational, we might capture something other than the causal effect of apologies. For example, individuals with lower contribution ranks might tend to apologise and raise their contributions as well, which leads to an effect of apologies. Our regression specification might not have done enough to control for such confounds. Running the same analysis with simulated apologies in the control treatment and finding no effects of apologies would help allay such concerns.

\textsuperscript{29}The results are the qualitatively similar if we simulate apologies in the control based on apology choices in the private and public treatment separately, see Tables A.1.5, A.1.6, A.1.7 and A.1.8.
Likewise, the placebo tests also indicate that for “synthetic apologies”, norms of apology in groups have minimal effects. In a regression of average group contributions on (simulated) group norms of sincerity and admitting error, $\beta_{\text{sincere}} = 3.750 \ (p = 0.558)$, $\beta_{\text{admit error}} = 6.444 \ (p = 0.531)$. There is also no indication of punitive action when norms of apology are not followed. Table 1.12 shows that contribution beliefs are not less optimistic when “synthetic” apologies are not used to admit errors or when they are not sincere. Furthermore, there is no evidence of a backfire effect when apologising with no intention to make amends, see Table A.1.4.

Table 1.11: Effects of “Apologies” in Control: Beliefs

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Contribution Belief</th>
<th>Niceness Belief</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apology Sent</td>
<td>-0.182</td>
<td>-0.036</td>
</tr>
<tr>
<td></td>
<td>[0.418]</td>
<td>[0.436]</td>
</tr>
<tr>
<td>Apology Received</td>
<td>-0.146</td>
<td>-0.018</td>
</tr>
<tr>
<td></td>
<td>[0.439]</td>
<td>[0.496]</td>
</tr>
<tr>
<td>No. of 3rd Party Apologies Received by Partner</td>
<td>-0.111</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>[0.471]</td>
<td>[0.546]</td>
</tr>
<tr>
<td>Constant</td>
<td>-3.497**</td>
<td>1.150***</td>
</tr>
<tr>
<td></td>
<td>[0.015]</td>
<td>[0.000]</td>
</tr>
<tr>
<td>Past Round Controls</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Session, Round Fixed Effects</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>3420</td>
<td>3420</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Tobit regressions with individual-partner-round level observations. Standard errors clustered at the group level. P-values in square parentheses. Estimates are based on the means of 2000 sets of apology simulations in the control group. Controls: Past contribution of partner, Past contribution of others besides partner, Past contribution rank, Partner’s past contribution rank, Past coordination, Past belief. Apologies are restricted to those made at non-highest rank.

Table 1.12: Effects of Partner’s Usage of “Apologies” on Beliefs in Control

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Contribution Belief</th>
<th>Niceness Belief</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partner’s Past Sincerity</td>
<td>0.610</td>
<td>-0.098</td>
</tr>
<tr>
<td></td>
<td>[0.217]</td>
<td>[0.402]</td>
</tr>
<tr>
<td>Partner’s Past Admission of Errors</td>
<td>-0.282</td>
<td>0.063</td>
</tr>
<tr>
<td></td>
<td>[0.445]</td>
<td>[0.432]</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.235</td>
<td>0.907***</td>
</tr>
<tr>
<td></td>
<td>[0.142]</td>
<td>[0.008]</td>
</tr>
<tr>
<td>Past Round Controls</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Session, Round Fixed Effects</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>2432</td>
<td>2432</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Tobit regressions with individual-partner-round level observations. Standard errors clustered at the group level. P-values in square parentheses. Estimates are based on the means of 2000 sets of apology simulations in the control group. Past Admission of Errors is the proportion of rounds in the past where a partner apologises when contributing strictly less. Past Sincerity is the proportion of rounds in the past where a partner strictly increases contributions after apologising. Controls: Past contribution of partner, Past contribution of others besides partner, Past contribution rank, Partner’s past contribution rank, Past coordination, Past belief.

Table 1.13 also shows no effects of group norms of sincerity on the evolution of cooperation and beliefs. Although this may be due to apologies being simulated based on
pooled apology choices in the private and public treatments, Table A.1.8 shows that there are no qualitative differences when apologies are simulated based on the private and public treatment separately.

Table 1.13: Effects of Group Usage of “Apologies” in Control

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Contribution</th>
<th>Niceness</th>
<th>Contributions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Belief</td>
<td>Belief</td>
<td></td>
</tr>
<tr>
<td>Group Norm of Sincerity</td>
<td>1.455</td>
<td>0.264</td>
<td>4.208</td>
</tr>
<tr>
<td></td>
<td>[0.182]</td>
<td>[0.394]</td>
<td>[0.174]</td>
</tr>
<tr>
<td>Group Norm of Admitting Error</td>
<td>1.489</td>
<td>0.352</td>
<td>10.578</td>
</tr>
<tr>
<td></td>
<td>[0.385]</td>
<td>[0.295]</td>
<td>[0.195]</td>
</tr>
<tr>
<td>Constant</td>
<td>-4.705**</td>
<td>0.907</td>
<td>-8.545*</td>
</tr>
<tr>
<td></td>
<td>[0.025]</td>
<td>[0.161]</td>
<td>[0.099]</td>
</tr>
<tr>
<td>Past Round Controls and Apologies</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Session, Round Fixed Effects</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>3420</td>
<td>3420</td>
<td>1140</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Columns 1 - 2: Tobit regressions with individual-partner-round level observations. Column 3: Tobit regression with individual-round level observations. Standard errors clustered at the group level. P-values in square parentheses. Estimates are based on the means of 2000 sets of apology simulations in the control group. Apologies are restricted to those made at non-highest rank. Group norm of admitting error is measured as the proportion of times in the group where an apology is made in cases where a partner has a strictly higher contribution. Group norm of sincerity is measured as the proportion of times in the group where an apology is followed by a strict increase in contributions.

Controls (Columns 1 - 2): Past contribution of partner, Past contribution of others besides partner, Past contribution rank, Partner’s past contribution rank, Past coordination, Past belief. Standard apologies received and sent. Controls (Column 3): Past contribution, Avg of others’ past contributions, Past contribution rank, Past Coordination, Standard apologies sent.

The evidence from the above placebo tests largely suggests that our earlier findings are not spurious and it is the act of apologising which contributes to improving cooperation in the public and private treatments. Actions, i.e., like raising contributions, are by themselves not sufficient to improve cooperation; words of apology also play a very important role, even more so when they are publicly made.

1.5 Discussion

In this paper, we study how a positive communication act — apologies — affects cooperation (and consequently social welfare) in public goods games. While standard models of cheap talk predict no effect, we find that the opportunity to apologise significantly improves cooperation within a public goods game. Generally, this is because sending and receiving apologies has positive effects on contributions and beliefs about others.\footnote{Nevertheless, care should be taken in extrapolating these results, which concern apologies that are more informal in nature, to the context of apologies which are more formal in nature. One example would be the head of a company apologising to shareholders for corrupt activities undertaken by the organisation. Tavuchis (1991) highlights how the mechanisms behind such formal apologies may differ from interpersonal apologies.}
Broadly, our paper contributes to an understanding of how norms interact with communication in building social capital and sustaining cooperation. Apologies are more than just individual communication and involve specific norms — participants show an understanding of when one should apologise, what to do after apologising and what to expect from an apologiser. When apologisers are not sincere, recipients implicitly sanction them. These factors are likely the reason why a simple, free message of “I am sorry” is not just cheap talk.\textsuperscript{31} That Bochet et al. (2006) and Wilson and Sell (1997) which involve numerical signals have little to no effects corroborate our view that communication can be complex.

Our experiment also demonstrates the importance of common knowledge in bringing out the best of norms. Stronger cooperative norms are established in groups where the norms associated with apologies are followed — especially in the public treatment where such information is common knowledge. This is consistent with a literature which discusses using social information as nudges, see Bicchieri and Dimant, 2019 for a review.

Interestingly, despite the strong effects of apologies, individuals who do not experience it do not appreciate it that much. When asked if they wanted the opportunity to apologise, individuals in the control treatment were more likely to disagree, compared to individuals in the apology treatments who were more likely to agree.\textsuperscript{32} Civility is an important aspect of civilisation, as emphasised by the sociologist Norbert Elias (Elias et al., 2000). Our experiment highlights that these seemingly trivial aspects of social etiquette which we often take for granted, may in fact be part of the social fabric which weaves us together.

\textsuperscript{31}Here, the signalling effect of apologies appears to be more sustained by its social consequences rather than being associated with individuals’ dispositions or valuations of relationships (Ho, 2012; Ohtsubo et al., 2018; Ohtsubo and Watanabe, 2009; Ohtsubo et al., 2012). Using first round contributions as a proxy for their cooperative types, we find positive effects on making sincere apologies only in the private treatment, see Tables A.1.9 and A.1.10 — this is not consistent with the public treatment outperforming the private treatment. “Type signalling” is thus unlikely to be a complete explanation for our results. We further examine guilt and shame, following research in psychology which show that proneness to feeling guilt and shame, is positively associated with repair behaviour in the form of apologies, although shame may sometimes lead to avoidance (Chrdileli and Kasser, 2018; Dunlop et al., 2015; Howell et al., 2012). Tables A.1.9 and A.1.10 show that such effects are generally weak. Neither do we find any differences between genders in the tendency to apologise ($p = 0.953, 0.714$ for public and private respectively) and sincerely apologise ($p = 0.753, 0.709$ for public and private respectively).

\textsuperscript{32}On a scale from 0 to 10, individuals in the public (private) treatment give responses which are 4.34 points (3.56 points) higher than the control, see Table A.1.11, Column 1. This also holds after controlling for cooperation within the group, see Table A.1.11, Column 2.
2 Ingroup Bias with Multiple Identities: The Case of Religion and Attitudes towards Government Size

with Daniel Sgroi and Jonathan Yeo

2.1 Introduction

Do you consider yourself Democrat or Republican? Religious, agnostic or atheist? How about your views on climate change, abortion, gun control? Meeting someone who shares your views on any of these issues or more can help construct a shared sense of identity and with it a tendency towards ingroup sentiment (Lane, 2016; Li, 2020). Numerous studies have found that even quite tenuous forms of group identity can produce remarkable ingroup bias, with random allocations of a color or an arbitrary label often being sufficient (Chen and Li, 2009; Tajfel et al., 1971). To date, a large body of work in this area has been directed towards uni-dimensional identity: the idea that individuals often fixate on a single salient dimension which defines themselves, e.g., race, gender, political affiliation or nationality etc. Nobel prize-winning economist and philosopher, Amartya Sen argues that this tendency for “us vs them” thinking has led to much bloodshed and suffering (Sen, 2006).

However, in practice identity is much more nuanced. There are infinitely many dimensions on which identity can be defined and it is hard to imagine anyone who has a narrow uni-dimensional focus in general. Attend a conference on gender issues and you may feel that your gender identity has become more important; later you may have a discussion on recent political events which instead brings your political affiliation to the forefront. Research has shown that priming different dimensions of identity can have disparate influences on performance in quantitative/verbal tests, cooperation and economic choices (Benjamin et al., 2010; Chen et al., 2014; LeBoeuf et al., 2010; Ravetti et al., 2019; Shih et al., 1999, 2006).

When multiple dimensions of identity are present, things become even more complex. How does one juggle the many ways in which one is similar or different to others and how does this influence decisions to help or hinder others? The way individuals react to and manage these cross-cutting and polarising dimensions of identity would have broad implications for harmony in society (Putnam, 2007). These issues motivate our study of how additional dimensions of identity influence individuals’ behaviour and the mechanisms behind them. While there has been some research in social psychology, political science and economics examining how the social and economic factors behind identities affect their relative strength (Atkin et al., 2021; Berman et al., 2020;
Eifert et al., 2010; Roccas, 2003; Rubin et al., 2016), they do not directly manipulate the salience of different dimensions of identity. Hence, they do not speak to our research agenda of studying the psychological mechanisms behind effects of multiple identity dimensions.

More related to our work, is a literature on multiple categorisation which examines how multiple dimensions of social identities can have an effect on inter-group relations, see Prati et al. (2021) for a review. Studies have found that when many non-overlapping (exogenous) dimensions of identity are salient, the complexity of identities, by reducing simple social categorisation, can lower inter-group friction. In other contexts, effects of multiple identities are however more nuanced. Nevertheless, most of such work has focused on the cognitive aspects of categorisation and neglected the social aspects of categories. When dimensions of identity themselves have moral and political connotations, effects can become even more complex. Without a controlled study it is not clear how multiple dimensions of identities might interact with their social content.

In this pre-registered study (Sgroi et al., 2021), we experimentally investigate the impact of additional dimensions of identity on ingroup bias (or outgroup discrimination) in a context where identity dimensions are politically charged. We recruited a sample of 961 participants in America using the Prolific platform between 20th May and 9th June 2021. Based on responses to questions posed to participants, they were classified into groups to make their identities salient. Subsequently, participants played a series of third party dictator games where they were asked to allocate resources between two different partners of known identity in a single dimension. For example, if this dimension was attitudes towards government size, one partner might be in favor of small government, the other in favor of large government. These allocations were then used to measure the extent of ingroup bias. We compare using separate treatments, the extent of bias when identities consist of two salient dimensions, to when identities consist of only a single salient dimension. Our two dimensions of identity — based on opinions about 1) government size and 2) religion — were carefully selected in line with past research which finds that these are two factors of similar importance and largely orthogonal in the respective dimensions of economic and social conservatism/liberalism (Everett, 2013). These dimensions of identity also speak to the recent increase in political polarization in America (Dixit and Weibull, 2007; Fiorina and Abrams, 2008).

Our basic results mirror standard findings in that there are strong ingroup biases throughout — participants allocate more to those who share a common identity. However, there is some heterogeneity in the magnitudes of ingroup bias. When only a single dimension is salient, the amount of ingroup bias is significantly smaller for the religion
dimension compared to the government dimension. These low levels of ingroup bias on the religion dimension are present despite high levels of identity strength recorded in the post experiment survey. We refer to identity dimensions with mild levels of ingroup bias despite strong group identity as “harmonious”. Identity based on attitudes towards religion hence seem to be in a domain which is more harmonious than that based on attitudes towards government size.

Furthermore, we find that compared to when only a single dimension (either government or religion) is salient, making the second dimension salient as well also changes behaviour. Again, the domain of identities matters — our results suggest that the effect of adding a second dimension differs depending on how “harmonious” the initial single dimension was. When we consider allocations under the more harmonious dimension (religion), making a second less harmonious (or more discordant) dimension (government) salient results in a sharp increase in ingroup bias. In contrast, for allocations on the more discordant dimension (government), making the second dimension (religion) salient has muted effects. This is counter-intuitive as we might assume that adding a separate orthogonal dimension should never worsen ingroup biases in the first dimension. However, it seems that effects can go in the other direction, by changing the domain to a more discordant social context.

2.2 Methods

This study used an online experiment to examine the effect of the number of salient dimensions of identity on identity strength and ingroup bias. The experiment had two different treatments which were conducted between subjects. There were two main stages in the experiment. The first stage was an identity inducement stage split into two treatments. Participants were randomly assigned to a treatment in which either 1) one dimension of identity was made salient, or 2) two dimensions of identity were made salient. In the second stage, participants completed an incentivized third-party allocation task which was used to elicit their level of ingroup bias. At the end of the experiment, there was a post-experiment questionnaire which included standard demographic questions together with questions concerning feelings of closeness to their ingroup, and their beliefs about the distribution of identities on each dimension (which were incentivized). The full experimental script is provided in Appendix B.2.

2.2.1 Identity Inducement

In our experiment, we focused on two main dimensions of political identity inspired by Everett (2013): 1) economic and 2) social conservatism/liberalism. Participants were shown a general statement on limited government (“Government intervention is
good for society”) and/or religion (“The world was created by a divine entity”) and asked to choose where they stood on the issue (agree/disagree). After choosing their position, participants were then shown their implied group membership based on their choices, with graphic symbols for each group to further increase identity salience. In the one-dimension treatment, they were only asked about one (random) statement. In the two-dimension treatment, they were asked about both statements in a random order — this order was kept the same for any questions thereafter. For completeness, participants in the one dimension treatment were asked in the post-experiment survey to choose their stand on the other dimension.

Our focus on political identities was based on observations in an early pilot that people form strong identities around their political affiliation. Religion and government size were chosen because they had several useful characteristics as discussed by Everett (2013):

- They are documented to be quite divisive and as such there is potential for them to initiate identity formation. Religion and Limited Government are 2 items in the 12-item Social and Economic Conservatism Scale (SECS) which weigh heavily in the “Social Conservatism” and “Economic Conservatism” factors (see Table 4 in Everett (2013)).

- They are of similar “strength” to avoid any asymmetry between dimensions complicating the effect of making another identity dimension salient. Religion and Limited Government have high factor loadings (> 0.8) in social and economic conservatism respectively (see Table 4 in Everett (2013)).

- They are not strongly correlated which preserves the notion that both dimensions are relatively independent. In the paper, the two factors had a weak correlation of 0.09 which was not significant at the 0.05 level (see Table 1 in Everett (2013)).

We also note that there is no obvious ordering in terms of status or hierarchy across different opinions in these two dimensions.

### 2.2.2 Third-Party Allocation Task

We used a third-party allocation task to elicit participants’ levels of ingroup bias in an incentivized way. Participants had to decide how they would allocate 100 experimental credits between two randomly selected other participants in the following manner. Firstly, they decided for a given dimension of identity (i.e., religion or government), how to allocate the credits in 3 different cases: 1) if one of the two selected participants has an ingroup identity while the other has an outgroup identity, 2) both have an ingroup identity and 3) both have an outgroup identity. In the two-dimension treat-
ment, the fact that on the other dimension of identity, the participant might have an ingroup or outgroup identity was highlighted.

Secondly, each participant had to decide whether they wanted to implement the allocation based on their aforementioned decisions. If not, they could instead elect to equally divide the experimental credits or randomly divide the credits between the two participants. For those in the two-dimension treatment, they had to make the same set of decisions on the second dimension. Each participant’s decisions were then used to determine the payoffs of two randomly selected other participants. For those in the two-dimension treatments, one set of decisions was randomly selected to be implemented.

From this task, we can measure their ingroup bias by calculating: $\text{Allocation}_{\text{ingroup}} - \text{Allocation}_{\text{outgroup}}$. This can be done for both the raw allocation (before they made the second decision), and the effective allocation (after they made the second decision). In our main analysis, we use effective allocations to calculate ingroup bias as these were the final decisions that were implemented for payment. Results are qualitatively similar and in fact quantitatively larger if raw allocations are used to calculate ingroup bias. See Tables B.1.4 - B.1.6 in Appendix for the corresponding results using raw ingroup bias.

### 2.2.3 Post-experiment Questionnaire

In the post-experiment questionnaire, we collected standard information on participants’ demographic characteristics like their birth year, gender, education level, personal income level and ethnicity etc. Furthermore, we also had questions relating to the task like their understanding of the instructions as well as how they made their allocation decisions. These were used in robustness checks of our main analysis. All results go through qualitatively, controlling for these variables in similar regressions, see Tables B.1.7 - B.1.10 in Appendix. We also had several questions which were used to construct variables used in exploring potential mechanisms behind our results.

Firstly, all participants were asked to give estimates of the unconditional conditional distributions of participants’ identities on the religion and government dimensions. In particular they were asked about (i) the percentage of all participants who belong to the group that agree/disagree with the religion statement, (ii) the percentage of all participants who belong to the group that agree/disagree with the government statement, (iii) the percentage of the participants who agree with the religion (government for session 3) statement that agree/disagree with the government (religion for session 3) statement, (iv) the percentage of the participants who disagree with the religion (gov-
ernment for session 3) statement that agree/disagree with the government (religion for session 3) statement. These estimates were elicited in an incentivized way: one of the four estimates was randomly selected for payment; the closer the estimate to the true value, the higher the payment.

Using these estimates and participants’ identities, we constructed participants’ ingroup correlation beliefs by calculating: \[ \text{Prob(Ingroup on second dimension | Ingroup on first dimension)} - \text{Prob(Ingroup on second dimension | Outgroup on first dimension)} \]. In cases where the first dimension does not correspond to the dimension that was conditioned on in questions (iii) and (iv), conditional probabilities had to be calculated via Bayes law and could fall outside of the possible range 0-100. These out-of-range conditional probabilities are treated as missing in the analysis. The distributions of calculated and directly observed conditional probabilities treating out-of-range values as missing are similar to the distributions of only directly observed conditional probabilities, k-smirnov test, \( p = 0.804 \).

Secondly, participants were asked to indicate on an 11-point Likert scale (0 to 10) to what extent they see themselves as belonging to their ingroup on religion and government dimension respectively. This question was adapted from the measure of group identification in Doosje et al. (2002), and is used to measure participants’ identity strength on each dimension.

### 2.2.4 Data Collection

The experiment was programmed in Otree (Chen et al., 2016) and conducted on the online research platform Prolific (www.prolific.co). Recruitment was restricted to participants currently residing in the United States and having a Prolific approval rate not lower than 95. Ethical approval for conducting the experiment was obtained from the University of Warwick Humanities and Social Sciences Research Ethics Committee (ref. HSSREC 178/19-20). We conducted 3 sessions during late May and early June in 2021. In total 480 + 240 + 241 = 961 participants completed the experiment in the 3 sessions respectively. Of this, 485 participants were in the one-dimension treatment, while 476 participants were in the two-dimension treatment. They were paid a completion fee of $0.9 together with a bonus fee based on decisions made by participants during the experiment. On average, participants spent 7 minutes in the experiment and received a payment of $1.41 which is above the recommended hourly payment rate of $9.60 for Prolific experimental participants.

Stratified randomisation was used to reduce sampling error. There were six strata given by a participant’s political and religious affiliation as indicated on Prolific’s demo-
graphic screener: Demographic/Republican/Independent × Religious/Non-religious. These two variables were chosen for stratification because they are closely related to the identities in the experiment and we suspected that behaviour in the experiment could differ by participant’s identities (this is indeed true, opinions on government and religion are significantly different across strata, Krusal Wallis test, $p < 0.001$ for both government and religion).

Simple random sampling was applied within each stratum: half of the participants in each stratum was randomly assigned to the one-dimension treatment, while the other half was assigned to the two-dimension treatment. We implemented an equal size for the one dimension and two dimension treatments because pilots indicated no significant differences in the standard deviation for ingroup allocations between them. Table B.1.3 compares demographic variables in the two treatments and show that they are relatively similar.

Our main analysis focuses on the measures of effective ingroup bias calculated from participants allocation decisions. Note that in the two-dimension treatment, each participant has a measure of ingroup bias for each dimension. In the one-dimension treatment, each participant has only one measure of ingroup bias for the dimension he/she faces. In total, we have $485 + 476 \times 2 = 1437$ data points.

## 2.3 Results

To ground our discussion, we first describe the distribution of identity groups in our study. Table 2.1 shows that there are relatively strong levels of polarisation on each of the religion and government dimensions. 58% of participants disagree with the statement “The world was created by a divine entity” while 31% disagree with the statement “Government intervention is good for society”. There is a negligible relationship between these two dimensions of identity, phi-correlation = -0.17.

<table>
<thead>
<tr>
<th>The world was created by a divine entity</th>
<th>Government intervention is good for society</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disagree</td>
<td>Disagree $n = 139$</td>
</tr>
<tr>
<td>Agree</td>
<td>$n = 167$</td>
</tr>
<tr>
<td>Total</td>
<td>$n = 306$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 2.3.1 Treatment Effects

Here, we analyse the aggregate treatment effects. In particular, we compare levels of ingroup bias when the two dimensions of identity (religion and government) are salient, to the alternative when only one dimension of identity (either religion or government) is salient. Ingroup bias is measured as the difference between allocations made to an ingroup member versus an outgroup member.

Figure 2.1 compares ingroup bias for the religion and government dimensions separately. Observe that in all cases, ingroup bias is significantly positive ($p < 0.001$ for all cases, one-sided sign test) — allocations towards ingroup members are significantly greater than that towards outgroup members. This is consistent with the existence of strong identity effects relating to religious belief and views about government size. In the post-experiment survey, we measured their identity strength by eliciting their feelings of closeness to their ingroup on an 11-point Likert scale (0 to 10), see Appendix B.2 for more details. On average, participants stated an average of 7.18 points which is significantly different from being neutral (5 points), one-sided sign test, $p < 0.001$. Their feelings of identity strength are also strongly correlated with allocations to their ingroup, see Table B.1.1.

![Figure 2.1: Effects of multiple salient dimensions on ingroup bias by identity dimension](image)

Plots of 95% confidence intervals of the mean in each treatment. Ingroup bias is the difference (in the percentage of resources) between allocations made to an ingroup versus an outgroup in a third-party dictator game.

However, we also note that there is some heterogeneity in ingroup bias across the cases. When only a single dimension of identity is salient, ingroup bias is significantly lower for allocations on the religion dimension compared to the government dimension ($p = 0.005$, two-sided Mann-Whitney test). When both dimensions of identity are salient, the amount of ingroup bias is however indistinguishable for allocations on the religion and government dimensions ($p = 0.803$, two-sided Mann-Whitney test). The effects of multiple dimensions on ingroup bias can be seen to be dependent on whether allocations involve the religion or government dimensions of identity. Mak-
ing a second dimension salient increases ingroup bias for allocations on the religion dimension, but has negligible effects for allocations on the government dimension.

Table 2.2 confirms the last observation using a multiple regression analysis with various controls. On the religious dimension, two salient dimensions of identity, compared to one salient dimension results in an increase in ingroup bias of 7.6 percentage points, Wald-test, $p = 0.012$. By contrast, on the government dimension, there is an insignificant decrease in ingroup bias of 0.12 percentage points, Wald-test, $p = 0.966$. This 7.7 percentage point difference is statistically significant, Wald-test, $p = 0.031$. There is thus strong evidence that the effects of multiple dimensions of identity can have disparate effects on ingroup bias in different social dimensions.

### Table 2.2: Marginal effects of two salient dimensions, relative to one salient dimension, on ingroup bias in each identity dimension

<table>
<thead>
<tr>
<th>Subgroup:</th>
<th>Religion Dimension</th>
<th>Government Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two Salient Dimensions</td>
<td>7.615** (3.035)</td>
<td>-0.118 (2.746)</td>
</tr>
</tbody>
</table>

Observations 1437

* 0.10 ** 0.05 *** 0.01. Estimates are from an ordinary least squares regression using individual-question level observations. It regresses ingroup bias on the interaction between dummies for the number of salient dimensions and dummies for the dimension of allocation. Includes controls for strata, session and order fixed effects. Standard errors clustered at individual level in parentheses.

#### 2.3.2 Possible Explanations

##### Differences in Identity Strength?

The treatment effects suggest that there are some fundamental differences between identity defined with respect to religion and government. One possible explanation might be that *stronger identities on the government dimension* have spillover effects on the religion dimension, raising the importance of religious differences and consequently increasing ingroup bias. There are several pieces of evidence which suggest that this is *not* the case.

Firstly, questions from our post-experiment survey indicate that the strength of identities on the religious dimension is greater compared to the government dimension. Table 2.3, Column 1 shows that on a scale from 0 to 10, participants’ feelings of belonging to their government ingroups were on average 0.57 points lower compared to religious ingroups, Wald-test, $p < 0.001$. Second, there is no evidence that making
the government dimension salient increases feelings of belong to one’s religious ingroup. Table 2.3, Column 2 shows that there is no significant effect of the number of dimensions on feelings of belonging to one’s religious ingroup, Wald-test, $p = 0.688$, nor one’s government ingroup, Wald-test, $p = 0.832$. Thus, it is unlikely that the rise in ingroup bias in the religion dimension when multiple identities are salient can be explained via the aforementioned mechanism.

Table 2.3: Effects of multiple salient dimensions by identity dimension on survey-measured feelings of belonging

<table>
<thead>
<tr>
<th>Dep Var: Feelings of belonging to ingroup</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government dimension</td>
<td>-0.568***</td>
<td>-0.654***</td>
</tr>
<tr>
<td></td>
<td>(0.121)</td>
<td>(0.210)</td>
</tr>
<tr>
<td>2 salient dimensions × Religion dimension</td>
<td>-0.091</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.226)</td>
<td></td>
</tr>
<tr>
<td>2 salient dimensions × Government dimension</td>
<td>0.038</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.180)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>7.466***</td>
<td>7.527***</td>
</tr>
<tr>
<td></td>
<td>(0.106)</td>
<td>(0.175)</td>
</tr>
<tr>
<td>Strata, Session and Order Fixed effects</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>1437</td>
<td>1437</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Ordinary least squares regressions with individual-question level observations. Standard error in parentheses clustered at the individual level.

**Signalling**

A second possible explanation might be that the higher levels of ingroup bias on the government dimension spillover to the religion dimension when both dimensions are salient because of a *signalling effect*. In particular, individuals might see membership of one group as a possible signal that their partner may also be a member of another group — especially when multiple dimensions of identity are salient. In our experiment, this might explain our results if an individual believes a partner who shares a group identity on the religious dimension is more likely to also have a shared identity on the government dimension when both dimensions are salient. Since the government dimension has higher levels of ingroup bias, such beliefs would lead to higher levels of ingroup bias on the religion dimension — even if ingroup bias on the single salient religion dimension is small.

We are able to examine such an explanation as in the post-experiment survey, participants were asked about their beliefs about the chance that ingroup/outgroup member on the religion (government) dimension is also an ingroup/outgroup member on the government (religion) dimension, see Appendix B.2 for more details. Using this, we
can calculate for each dimension, in percentage points, the extent to which participants are *more likely* to believe that an ingroup member on a particular dimension is also an ingroup member on the other dimension (relative to an outgroup member on the particular dimension). For exposition purposes, we refer to this as their *ingroup correlation beliefs*.

Table 2.4: Marginal effects of two salient dimensions relative to one salient dimension, and ingroup correlation beliefs, on ingroup bias in each identity dimension

<table>
<thead>
<tr>
<th>Subgroup:</th>
<th>Religion Dimension</th>
<th>Government Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 salient dimensions</td>
<td>8.050***</td>
<td>-3.979</td>
</tr>
<tr>
<td>(3.052)</td>
<td>(3.152)</td>
<td></td>
</tr>
<tr>
<td>Ingroup correlation beliefs</td>
<td>0.188***</td>
<td>-0.001</td>
</tr>
<tr>
<td>(0.048)</td>
<td>(0.044)</td>
<td></td>
</tr>
</tbody>
</table>

Observations 1232

| * 0.10 ** 0.05 *** 0.01. Estimates are from an ordinary least squares regression using individual-question level observations. It regresses ingroup bias on the interaction between dummies for the number of salient dimensions, dummies for the dimension of allocation and ingroup correlation beliefs. 215 observations excluded due to out of range beliefs which are treated as missing. Includes controls for strata, session and order fixed effects. Standard errors clustered at individual level in parentheses. |

Results of our analyses are inconsistent with the above signalling explanation. Firstly, while Table B.1.2 shows that on average, participants have positive ingroup correlation beliefs, there are limited effects of making two dimensions of identity salient. Secondly, controlling for their ingroup correlation beliefs in regressions similar to Table 2.2 does not explain away our earlier results on the heterogeneous impacts of multiple dimensions of identity on ingroup bias; in fact differences are even bigger. Table 2.4 shows that while ingroup correlation beliefs have some explanatory effect on ingroup bias on the religion dimension, even after controlling for it, the effect of two salient dimensions on ingroup bias is still 12 percentage points higher on the religion dimension compared to government dimension, Wald-test, *p* = 0.002.

**Identities in conflict vs Identities in Harmony**

In the previous section, we showed that differences in identity strength or signalling cannot explain our treatment effects. We next consider how another facet of identities — the degree of “conflict” or “harmony” between identity groups defined on specific dimensions — may explain our results. While opinions on government and religion
are polarized, we first note an interesting difference between sentiment directed towards those who do not share identity with respect to religion as opposed to government. Table 2.5 examines the marginal effect of identity strength on ingroup bias in each identity dimension. A 1 point higher answer on the 11-point Likert scale for identity strength is associated with a 4.3 percentage point increase in ingroup bias on the government dimension, but only a 2.2 percentage point increase in ingroup bias on the religion dimension. The difference is statistically significant, Wald-test, \( p = 0.021 \). That identity strength matters much less for allocation decisions on the religion dimension suggests that there is a greater sense of harmony between individuals of different religious opinion, as compared to government opinion.

Table 2.5: Marginal effects of feelings of belonging to ingroup, on ingroup bias in each identity dimension

<table>
<thead>
<tr>
<th>Subgroup:</th>
<th>Religion Dimension</th>
<th>Government Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marginal Effect on ingroup bias of:</td>
<td>2.223***</td>
<td>4.308***</td>
</tr>
<tr>
<td>Feelings of belonging to ingroup</td>
<td>(0.593)</td>
<td>(0.732)</td>
</tr>
<tr>
<td>Observations</td>
<td>1437</td>
<td></td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Estimates are from an ordinary least squares regression using individual-question level observations. It regresses ingroup bias on the interaction between feelings of belonging to ingroup and dummies for the dimension of allocation. Includes controls for strata, session and order fixed effects. Standard errors clustered at individual level in parentheses.

A greater sense of harmony between groups should be associated with higher concerns of fairness and mutual respect. We have several pieces of evidence that illustrate this. Using an exact 50-50 split between ingroups and outgroup as a proxy for concerns of fairness and mutual respect, Figure 2.2 shows that for decisions in the 1 dimension treatment, participants show much higher concerns for fairness in the religion dimension. When making allocations on the religion dimension, 75 percent are an equal split. In contrast, when making allocations on the government dimension, only 64 percent are an equal split. This difference is statistically significant, two-sample test of proportions, \( p = 0.007 \).

Evidence from our post-experiment survey is also consistent with the above. Participants were given a multiple choice question on how they decided on their allocations. In the first two sessions, there was no explicit option to choose fairness, but if they selected “other”, it was possible to indicate fairness or mutual respect as a reason. In the third session, we added in fairness as an explicit option. In the one-dimension treatment, 38 out of 237 either selected fairness as their reason or gave an open-ended answer involving fairness or mutual respect for the religion dimension compared to 23
out of 248 for the government dimension. The difference in proportions of participants who gave fairness or mutual respect as a reason is significantly different, two-sample test of proportions, \( p = 0.025 \).

However, when making decisions in the two-dimension treatment, it seems that these concerns for fairness evaporate. Figure 2.2 shows that participants are almost as likely to choose an equal split on the religious dimension compared to the government dimension — there is no significant difference in the proportion of equal splits for the religion and government dimensions in the two dimension treatment, two-sample test of proportions, \( p = 0.166 \). Likewise, in the post-experiment survey, participants are equally likely to indicate fairness or mutual respect as a reason for allocations in the two dimension treatment. 30 out of 239 for the religion dimension, and 27 out of 237 for the government dimensions indicate mutual respect and fairness as a reason for their allocation. This difference is not significant, two-sample test of proportions, \( p = 0.697 \).

The above results suggest that making the relatively discordant government dimension salient has spillover effects on the relatively harmonious relations in the religion dimension. While interactions on the religion dimension are usually fairly harmonious, the knowledge that others might differ on the government dimension (for which there is greater conflict) crowds out any considerations of fairness or mutual respect that apply to the religion dimension. Interestingly, this spillover effect is asymmetric — we do not observe similar spillover effects of the religion dimension on the government dimension. Ingroup bias on the government dimensions and measures of concerns for fairness/mutual respect remain relatively unchanged when the religion dimension is made salient as well, see Figures 2.1 and 2.2.
2.4 Discussion

In our paper, we find evidence that raising the salience of additional dimensions of identity can have effects on ingroup bias — inattention to other possible dimensions of identity is thus likely important. However, our results are more nuanced compared to related research. In contrast to research on multiple categorisation where non-overlapping categories reduce ingroup bias, we find an increase in ingroup bias based on religion when multiple dimensions of identity are salient — this occurs despite the multiple dimensions being relatively uncorrelated.

Possible explanations for the failure of multiple categorisation in the literature include the causal centrality and dominance of particular identities (Chen and Urminsky, 2019) and the conceptual relation of categories (Crisp and Beck, 2005). However, exploring the potential mechanisms of differences in identity strength and signalling, we do not find any evidence consistent with them. One reason could be that our study involves dimensions which are more politically charged, involving elements of what columnist Erza Klein calls issue and identity based polarisation (Klein, 2020).

We find evidence that there is even more nuance amongst such (strong) polarised identities. In particular, it is possible to classify identities into different “domains” based on the kinds of intergroup relations. In our experiment, the religion identity dimension involves norms of fairness and mutual respect which are seemingly less present in the government identity dimension, perhaps due to social norms which suppress religious discrimination. Other studies have found evidence that on ethnic identity dimensions, similar social norms which suppress racial discrimination also seem to be present (Iyengar and Westwood, 2015).

Consequently, there appear to be different kinds of interactions between these domains when multiple dimensions of identity are present. Results show that the government dimension has negative spillover effects on the religion dimension reducing the apparent “harmony” that exists in our sample between those with or without religious identity (which exhibits itself as a low level of ingroup bias in allocations). However, the religion dimension does not have positive spillover effects on the government dimension, proving incapable of generating harmony where it did not already have a foothold. We surmise that this may be because salience of identities on the government dimension “activates” thinking on political partisanship lines. Such asymmetry also highlights the fragility of any notion of fairness and harmony between different identity groups.

Broadly, our results suggest that when studying multiple dimensions of identity, it is also important to consider the social context of dimensions of identity rather than
just attempting to draw general conclusions. We believe that more work needs to be done in linking the sociology of inter-group interactions to the psychology of multiple identities.
3 Mindfulness Reduces Information Avoidance

with Elliott Ash, Daniel Sgroi and Anthony Tuckwell

3.1 Introduction

A well-known bias in individual decision-making is the tendency to avoid information about potentially negative outcomes, even if it is freely available. Information avoidance can be costly: an individual’s ability to make good decisions hinges critically on their knowledge of the state of the world. Its potential costs are greater still when taking into account externalities; for example, individuals unwilling to learn about whether or not they carry an infectious disease pose a significant risk to society, as they may infect others.

Previous work suggests that anticipatory emotions (such as worry or regret) play an important role in information avoidance\(^1\). It is therefore plausible that mental training that targets the regulation of such emotions might help to diminish their influence on decision making. One such form of mental training is “mindfulness” meditation: a secularised form of Buddhist meditation, initially developed for pain management (Kabat-Zinn, 1990). Reporting evidence from a Randomised-Controlled Trial (RCT), this paper will examine whether mindfulness can influence information avoidance.

Mindfulness has become increasingly popular in the West in recent decades and has been linked with a variety of benefits, e.g. for health, stress, depression, and productivity (Brown et al., 2007). Meditation encourages a particular state of mind (non-judgmental attention to the present moment), and various evidence from psychology and neuroscience has demonstrated that its practice can increase levels of attention and emotion regulation (and, indeed, structurally change regions of the brain associated with such tasks\(^2\)). However, mindfulness can be viewed as a trait as well as a meditation practice (Brown and Ryan, 2003): different individuals naturally spend more or less time in such mindful states even if they have never meditated, so its study has implications for non-meditators as well.

To test for a causal effect of mindfulness on information avoidance, we designed a randomized human-subjects experiment. Experiment participants \((n = 261)\) were randomly allocated to either a treatment intervention (14 days of 15-minute guided mindfulness meditations), or an active control intervention (14 days of 15-minute guided

---

\(^1\)See Golman et al. (2017) for a review.

\(^2\)See Hölzel et al. (2011) for a review.
relaxing-music listening. This design allows us to test the effects of mindfulness over-and-above just feeling more relaxed.

Our main finding is that mindfulness reduced information avoidance – that is, the tendency for a participant to avoid receiving information that might cause worry or regret (Ho et al., 2020). Relative to the active control, mindfulness treatment reduced score on the information avoidance scale by approximately 0.25 standard deviations. We provide supporting evidence for emotion regulation as a plausible mechanism, as the treatment had a positive effect on a self-report measure of non-reactivity to inner experience.

We next review the relevant literature and our relative contribution, before detailing the design and results. We end with a discussion and some concluding remarks.

### 3.2 Literature

This research adds to the literature on the causes of information avoidance. The previous literature has documented various potential causes of information avoidance, with Golman et al. (2017) grouping them into hedonic (avoiding information to avoid feeling bad, e.g. because of belief-based utility) and strategic (as a way to committing to an a priori preferred course of action). However, relatively little is understood about the psychological and cognitive forces that make different individuals more or less susceptible to avoiding information. In part this could be because of a lack of a measure of information avoidance as a psychological construct, which was the motivation for Ho et al. (2020) to produce the scale that we use in this paper. Sweeny et al. (2010) mention some empirical work that suggests coping styles and uncertainty orientation as two possible explanations for individual differences in information avoidance. Our paper adds to this literature by documenting the role of mindfulness.

We expect mindfulness to act on the hedonic form of information avoidance – where individuals avoid information that could potentially affect their beliefs because of psychological costs such as worry, regret, disappointment, pessimism or cognitive dissonance (Golman et al., 2017). The mindful state encourages individuals not to be wrapped up in thoughts and beliefs as if they were strictly true (the quality of “non-judgment”), and instead hold them lightly in awareness (a concept known as “meta-awareness”). Thus, mindfulness may weaken the potential emotional imprint of beliefs, reducing the influence of worry, regret, and other negative cognitive factors. In support of this idea, Saunders et al. (2013) find that mindfulness increases recall of self-threatening information. More indirectly, mindfulness has been shown to reduce

---

3 The same instructor delivered both the treatment and the active control.

4 See Schooler et al. (2011) for a review.
symptoms of belief-based utility, such as anxiety (Roemer et al., 2009) and habitual worrying (Verplanken and Fisher, 2014). In general, mindfulness has been found to increase ability to regulate emotions; for example, reducing emotional interference when performing a task (Ortner et al., 2007) and decreasing emotional reactivity (Goleman and Schwartz, 1976). Researchers point to people in mindful states being better able to “reappraise” emotions (Garland et al., 2011), which means they are more equipped to process uncomfortable emotions, and less likely to engage in experiential avoidance of thoughts and feelings. (Kumar et al., 2008). Supporting this work is neuroscience evidence showing that meditators have increased activation in regions of the brain associated with emotion regulation (Hölzel et al., 2011).

Our paper also relates to a literature that investigates the influence of mindfulness on decision-making. Alem et al. (2016) conduct an RCT which tested whether mindfulness influenced risk preferences, time preferences and health-related behaviours (e.g. smoking, eating, alcohol consumption, sleeping), but their results in general were not statistically significant. Moreover, their active control (watching a historical documentary) does not specifically control for being relaxed, so it is hard to disentangle the effects of being mindful from being relaxed in their results. The RCT in Noone and Hogan (2018) investigates the effects of mindfulness on various cognitive tasks (including a heuristics-and-biases measure), using the Headspace app as a mindfulness intervention and a sham meditation active control. They did not find statistically significant effects, but that could be due to a short treatment not being effective enough, or it could be that the sham meditation might have engendered some degree of mindfulness. Papers analyzing effects of mindfulness on decision-making have found that mindfulness can make decisions more adaptive (in a gambling context) (Lakey et al., 2007); reduce negativity bias (Kiken and Shook, 2011); reduce correspondence bias (Hopthrow et al., 2017); decrease the sunk cost effect (Hafenbrack et al., 2014); alleviate addiction and self-control problems; and increase levels of altruism (Iwamoto et al., 2020). Our paper adds information avoidance to these documented effects.

3.3 Experimental Design

3.3.1 Sample Recruitment

We recruited subjects using Prolific (www.prolific.co), an online crowd-sourcing platform (based in the UK) which connects researchers to participants for academic studies. Like the more commonly used platform MTurk, Prolific has been found to produce data of a comparable quality to more traditional participant pools (Peer et al., 2017) and has been used to successfully run experiments in economics (e.g. Marreiros et al.,

---

5See Zgierska et al. (2009) for a review
2017) and psychology (e.g. Callan et al., 2017). However, Prolific has the advantage of having participants who are more naive with respect to experimental tasks and less dishonest than those on MTurk (Peer et al., 2017). Another reason why we chose Prolific is that it is more active than MTurk in the UK, where our recruitment targeted.

We recruited 261 subjects in one wave. We restricted participation to UK residents to maximise comprehension and familiarity with the instructor’s English accent. Besides, we required that participants have completed at least 10 Prolific studies previously, with a a good participation track record (Prolific approval rate not lower than 95% ). We also pre-screened on meditation experience, recruiting only participants who had answered “No” to Prolific’s own pre-screening question, “Do you meditate?”. In the invitation to potential participants, the study was described as investigating the effects of mood on decision-making, and participation would involve doing a simple and enjoyable activity for 15 minutes a day on 14 consecutive days. To minimise attrition, participants were told on sign-up that their submissions would only be “accepted” (i.e. they would only be paid) if they completed all parts of the study (unless there were exceptional circumstances).

3.3.2 Interventions

After signing-up to the study and completing the pre-course survey, participants were randomly allocated to one of two groups: a mindfulness intervention (treatment) group, and a music intervention (active control) group. On each of the following 14 days, the instructions for the intervention activity were given by a professional instructor via an audio recording. On the day after the course, the participants took a survey which measured our outcomes. The software o-Tree (Chen et al., 2016) was used to host the surveys, while Qualtrics (www.qualtrics.com) was used to deliver the interventions.

**Mindfulness intervention.** Here the instructor led the participants in a guided mindfulness meditation each day. Each session started with a short introduction (welcoming the participants). The instructor then led the participants through three stages of meditation: (1) bringing awareness to now (noticing what is happening outside and how you are); (2) mindful breathing (being aware of the breath and cultivating an attitude of non-judgment as thoughts arise); and (3) a body scan (expanding this awareness from the breath to the entire body). This was then followed by a period where the participants were asked to just sit with whatever awareness they had accumulated, before the instructor came back to end the session.

---

6Calculations suggested a sample of 220-260 subjects would be adequate to detect effects with 80% power and 5% significance (Ash et al., 2020).
Music intervention. Here the same instructor led the participants in a period of relaxing music listening each day. The idea of the intervention was to try to control for as many of the structural elements of the treatment as possible (15 minutes a day of doing an activity instructed by an audio recording, with the same instructor leading the activity), and in addition control for the relaxing effects of the meditations. To try to make the instructor’s presence felt as much as in the treatment, the instructor spent time on a short introduction before the music began (welcoming the participants, mentioning the details of the artist/album etc., and also reciting a famous quote about music for the participants to contemplate), and after the music finished he came back to end the session.

In order to boost feelings of instructor-participant interaction for both groups (and help minimise attrition), the instructor prepared three short videos of himself to be played at the start, middle and end of the 14-day intervention period (simple check-ins). In addition, participants were sent daily reminders on Prolific about the activity sessions. Compliance was encouraged before the recordings began with a request to close all sources of distraction and to stay on the browser tab (and not multitask). Compliance was then monitored using two different measures: (1) how often they left their browser tab during the recording; (2) whether they clicked to the “next page” right at the moment when the instructor asked them to at the end of the recording. We also included an optional feedback question about their experience of the session at the end.

3.3.3 Procedure

The study was launched on August 27th, 2020. On the first day we recruited 261 participants, who signed up and completed the pre-course survey. Then from the 28th of August through to the 11th of September, each day the participants were invited to complete a session of the daily activity (study available from 6am; reminder sent at 3pm), and were asked to submit by 3am the following day. Participants who missed a session were asked to take the session on the following day instead. Participants who attempted a session but had difficulties finishing it for some reason (e.g. because of internet trouble) were allowed to miss the session. Any participant who missed more than one session without giving a reason was excluded. On the 12th of September, participants were asked to take the post-course survey.

Various studies document the salutary effects of music for stress – see de Witte et al. (2020) for an overview. In some contexts music has been found to have comparable effects to meditation in reducing stress (e.g. Innes et al., 2016), and has previously been used as part of an active control for the widely-used Mindfulness-Based Stress Reduction (MBSR) programme (MacCoon et al., 2012). Stress impacts cognitive processes (e.g. “System 1” and “System 2” thinking (Kahneman, 2011)) that underlie various kinds of decisions (including information avoidance).
The participants were paid for doing the activities (£2 per session in the first week; £2.50 per session in the second week) and taking the surveys (£2 for the pre-course survey; £3 for the post-course survey).

3.3.4 Outcomes

**Information avoidance.** We used the *Information Preference Scale* (IPS) (Ho et al., 2020): a 13-item scale (validated by an incentivised experiment) that measures an individual’s willingness to receive information that might cause worry or regret in a series of thirteen hypothetical scenarios.\(^8\) Participants indicate their responses to the scenarios using a 4-point scale coded \([0, \ldots, 3]\), giving scores \([0, \ldots, 39]\). To avoid revealing the purpose of the study, information preferences were measured in the post-course survey only.

**Mindfulness.** We used the 15-item version of the *Five Facet Mindfulness Questionnaire* (FFMQ) (Baer et al., 2012), a frequently-used measure of mindfulness and its underlying dimensions (Sauer et al., 2013).\(^9\) Responses are made on a 5-point scale coded \([0, \ldots, 4]\), giving a mindfulness score of \([0, \ldots, 60]\), but the scale can also be disaggregated into subscales that measure five attributes of mindfulness: observing, describing, acting with awareness, non-judging of inner experience, and non-reactivity to inner experience (3 items in each, scores \([0, \ldots, 12]\)). To avoid revealing the purpose of the study, this outcome was also measured in the post-course survey only.

**Stress.** We used the 10-item version of the *Perceived Stress Scale* (PSS) (Cohen and Williamson, 1988), a widely-used instrument to assess subjective perceptions of stress (Liu et al., 2020).\(^{10}\) Responses are made on a 5-point scale coded \([0, \ldots, 4]\), giving scores \([0, \ldots, 40]\). This outcome was measured in the pre-course survey and the post-course survey.

3.3.5 Regression Specification

To estimate the statistical effect of the treatment on outcomes in the post-course survey, we use the following linear regression model:

\[
Y_i = \alpha + \beta \text{Treat}_i + \gamma X_i + \epsilon_i \tag{3.1}
\]

where \(Y_i\) is the outcome, \(\text{Treat}_i\) is a dummy variable equal to 1 for individuals in the mindfulness treatment, and \(X_i\) is a vector of individual characteristics measured at

---

\(^8\)See Appendix C.1 for the items of the IPS.

\(^9\)See Appendix C.2 for the items of the FFMQ.

\(^{10}\)See Appendix C.3 for the items of the PSS.
3.4 Results

3.4.1 Sample Characteristics

Considering the sample characteristics, there were no significant differences in the means of our baseline measures across the treatment and control groups (see Appendix Table C.4.1). Levels of attrition were 13% in the treatment and 18% in the control, which mostly occurred after the pre-course survey (see Appendix Figure C.4.1); the remaining samples of non-attritors were still comparable on the baseline measures (see Appendix Table C.4.2). Average feedback of the sessions was positive and similar for the treatment and control (see Appendix Figure C.4.2), and there were strong and similar levels of compliance in both groups according to our different compliance measures (see Appendix Figure C.4.3 and Figure C.4.4).

3.4.2 Effects on Perceived Stress

The treatment intervention and active control intervention had similar effects on measured stress. As shown in Figure 3.1, both interventions reduced perceived stress, but the effects are not statistically significant. It is unclear why the treatment and active control did not have significant effects on reducing stress. It could be that there was an increasing time trend of stress which offset the effects of interventions. It could also be that the length of the interventions and amount of practice per day were insufficient to generate significant reductions, or the perceived stress scale was too noisy a measure to have detected a change with the current sample. In any case, the active control has fulfilled its primary purpose: to provide an equivalent effect on stress – neither pre-course nor post-course perceived stress level is significantly different in the control and the treatment group (Mann-Whitney test, $p = 0.669$ for pre-course, $p = 0.912$ for post-course).

3.4.3 Mindfulness and Information Avoidance

We now evaluate the effect of the treatment on information avoidance. As seen in Table 3.1, being assigned to the treatment had a significant positive effect on preferences to receive potentially negative information as measured by the Information Preference Scale (IPS) ($p = 0.060$ without demographics; $p = 0.047$ with them). Being in the treatment is associated with an increase of approximately 0.25 standard deviations in the information preference scale.
Figure 3.1: Intervention Effects on Perceived Stress

Notes: This figure shows the pre-course and post-course means of perceived stress in the treatment and control. Gray bars show 95% confidence intervals.

Table 3.1: Effect of the Treatment on Information Preferences

<table>
<thead>
<tr>
<th>Marginal effects</th>
<th>Information Preference Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Treatment</td>
<td>0.251*</td>
</tr>
<tr>
<td></td>
<td>(0.132)</td>
</tr>
<tr>
<td>Demographics</td>
<td>No</td>
</tr>
<tr>
<td>Observations</td>
<td>226</td>
</tr>
</tbody>
</table>

*p < 0.10. **p < 0.05. Marginal effects from OLS regressions with robust standard errors in parentheses. IPS is standardised. Demographics include sex, age, race, education, household income and conservatism. 2 observation excluded for column 2 due to non-reported age.
Table 3.2: Effect of the Treatment on Non-Reacting

<table>
<thead>
<tr>
<th>Marginal effects</th>
<th>Non-React Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Treatment</td>
<td>0.272**</td>
</tr>
<tr>
<td></td>
<td>(0.132)</td>
</tr>
<tr>
<td>Demographics</td>
<td>No</td>
</tr>
<tr>
<td>Observations</td>
<td>226</td>
</tr>
</tbody>
</table>

*p < 0.10, **p < 0.05. Marginal effects from OLS regressions with robust standard errors in parentheses. Non-React Scale is standardised. Demographics include sex, age, race, education, household income and conservatism. 2 observation excluded for column 2 due to non-reported age.

3.4.4 Emotion Regulation as a Potential Mechanism

In Table 3.2 we show that the treatment had a significant effect on the non-react scale of the FFMQ (p = 0.041 without demographics; p = 0.076 with them). In terms of magnitudes, being in the treatment group is associated with an increase of approximately 0.27 standard deviations in the non-react scale. The items of the scale are all about not reacting to distressing inner experience. This inner experience could include anticipatory emotions such as worry or regret. This evidence supports regulation of anticipatory emotions as a mechanism by which the mindfulness training was able to reduce tendencies for information avoidance.

3.5 Discussion

In this paper we have provided evidence on mindfulness as a cause of differences between individuals in their susceptibility to information avoidance. The costs of information avoidance for individuals, society and the economy are potentially substantial (from individuals unwilling to learn about their health, including whether or not they carry infectious diseases, to students unwilling to check their marks, to investors holding off looking at their stocks’ performance (Ho et al., 2020)) so understanding what might drive some individuals to avoid information more than others is important. Our evidence suggests that people in the population who spend more of their time inhabiting mindful states are better able to look at potentially negative, but nonetheless useful, information about themselves and the world. Supplementary evidence suggests that it may be mindfulness’s effects on emotion regulation (specifically, non-reaction to emotions) that acts as a potential mechanism through which this greater tolerance for information operates.

11The effects on the other facets of mindfulness were not as significant (see Appendix Figure C.4.5).
An important concern about the randomized control trial is whether subjects in the treatment group actually engaged with the guided meditations. The compliance measures were encouraging in this regard in that it appeared that significant proportions of the subjects were listening to the recordings (e.g. not switching off the browser tab, and clicking to the next page when the instructor asked them to at the end of the recording). However, it could be that the subjects listened to the recordings but did not practice the meditations. Although this is hard to rule out, it seems difficult to square with the evidence, which showed that subjects in the treatment group developed higher levels of non-reaction, a known effect of meditation. An additional concern is that subjects in the treatment group, once they knew that meditation was their daily activity, would have certain expectations about the effects of meditation, and this would then influence their responses on the information avoidance measure (an “experimenter demand” effect). Given that information avoidance is an unknown effect of meditation (not discussed in the public domain), and that no relevant cues were given during the interventions in relation to information avoidance, we are less concerned about experimenter demand in relation to this outcome. Nonetheless, we controlled the expectations that could be managed in the design as best as possible, with both the treatment and control groups being told the same message in regards to their activity at the start of the interventions: that it had been found to have a “positive effect on people’s mood and wellbeing”.

Our paper adds information avoidance to the growing list of documented benefits of mindfulness. This result has potentially strong policy implications. “Nudging” (Thaler and Sunstein, 2009) has become a staple of behavioural policy, being employed in various governments throughout the world. However, by shaping individual choices without their knowledge, it has been criticised as a potential threat to individual autonomy. Making better decisions through greater levels of mindfulness, on the other hand, is a fully conscious process, so mindfulness training could provide governments with a more ethical approach to ameliorating cognitive biases. Our evidence shows that mindfulness is able to reduce information avoidance, but more work is needed to test its effects on a wider array of cognitive biases. For example, mindfulness (by managing the emotions triggered by beliefs) might also affect the processes underlying “motivated beliefs” (such as wishful thinking). We hope our investigation will encourage more research in this area.

---

12 See, for example, Hausman and Welch (2010).
13 See Bénabou and Tirole (2016) for a review.
### Appendix A (for Chapter 1)

#### A.1 Appendix Figures and Tables

Figure A.1.1: Baseline traits in each treatment

![Figure A.1.1: Baseline traits in each treatment](image)

Table A.1.1: Apology Treatments and Persistent Cooperative Norms

<table>
<thead>
<tr>
<th></th>
<th>Last Round Contribution</th>
<th>Unconditional Cooperation</th>
<th>Conditional Cooperation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dep Var:</td>
<td>Contribution</td>
<td>Unconditional Cooperation</td>
</tr>
<tr>
<td>Private</td>
<td>1.394</td>
<td>1.446</td>
<td>-0.902</td>
</tr>
<tr>
<td></td>
<td>(1.289)</td>
<td>(0.959)</td>
<td>(0.829)</td>
</tr>
<tr>
<td>Public</td>
<td>2.678*</td>
<td>3.795***</td>
<td>1.707*</td>
</tr>
<tr>
<td></td>
<td>(1.489)</td>
<td>(1.045)</td>
<td>(0.992)</td>
</tr>
<tr>
<td>Constant</td>
<td>7.421***</td>
<td>7.254***</td>
<td>0.899</td>
</tr>
<tr>
<td></td>
<td>(1.947)</td>
<td>(1.616)</td>
<td>(0.928)</td>
</tr>
<tr>
<td>Session Fixed Effects</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>232</td>
<td>232</td>
<td>232</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. OLS regressions with individual level observations. Standard errors in parentheses clustered at the group level. Columns 2 to 4 report results from the elicitation of their cooperative types at the end of the experiment. Unconditional Cooperation is the amount they are willing to contribute, not knowing about other’s contributions during the type elicitation stage. Conditional Cooperation, constant and slope are the individual-level constants and slopes respectively, from regressions of how much they are willing to contribute, given different average contributions of others, using data from the strategy method in the type elicitation stage.
### Table A.1.2: Group Sincerity and Contribution Differences over time

<table>
<thead>
<tr>
<th>Dep Var: Contribution</th>
<th>Apology Treatments</th>
<th>Public Treatment</th>
<th>Private Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round=1 × Group Sincerity</td>
<td>-3.550 (2.389)</td>
<td>-4.675 (2.915)</td>
<td>-5.870 (4.506)</td>
</tr>
<tr>
<td>Round=2 × Group Sincerity</td>
<td>1.979 (2.292)</td>
<td>0.920 (3.731)</td>
<td>-1.957 (4.311)</td>
</tr>
<tr>
<td>Round=3 × Group Sincerity</td>
<td>5.409* (3.144)</td>
<td>7.000* (3.768)</td>
<td>-0.005 (4.341)</td>
</tr>
<tr>
<td>Round=4 × Group Sincerity</td>
<td>7.834*** (2.733)</td>
<td>11.765*** (3.768)</td>
<td>2.037 (4.052)</td>
</tr>
<tr>
<td>Round=5 × Group Sincerity</td>
<td>8.305*** (2.830)</td>
<td>14.288*** (3.815)</td>
<td>0.270 (4.082)</td>
</tr>
<tr>
<td>Round=7 × Group Sincerity</td>
<td>13.410*** (2.714)</td>
<td>13.942*** (1.932)</td>
<td>6.984 (4.315)</td>
</tr>
<tr>
<td>Round=8 × Group Sincerity</td>
<td>15.145*** (2.564)</td>
<td>15.193*** (2.461)</td>
<td>11.292** (4.765)</td>
</tr>
<tr>
<td>Round=10 × Group Sincerity</td>
<td>11.304*** (2.247)</td>
<td>8.421*** (2.753)</td>
<td>9.846** (4.519)</td>
</tr>
<tr>
<td>Round=12 × Group Sincerity</td>
<td>11.682*** (2.466)</td>
<td>13.313*** (2.999)</td>
<td>9.625* (4.973)</td>
</tr>
<tr>
<td>Round=14 × Group Sincerity</td>
<td>15.183*** (2.384)</td>
<td>17.033*** (2.264)</td>
<td>13.025** (4.868)</td>
</tr>
<tr>
<td>Round=16 × Group Sincerity</td>
<td>9.238*** (2.375)</td>
<td>15.376*** (4.674)</td>
<td>3.698 (4.877)</td>
</tr>
<tr>
<td>Constant</td>
<td>14.345*** (2.184)</td>
<td>17.874*** (2.378)</td>
<td>14.406** (5.183)</td>
</tr>
</tbody>
</table>

Session, Round Fixed Effects ✓ ✓ ✓
Group Norm of Admitting Error × ✓ ✓ ✓
Round Fixed Effects ✓

Observations 2496 1216 1280

* 0.10 ** 0.05 *** 0.01. OLS regressions with individual-round level observations. Standard errors in parentheses clustered at the group level.
Table A.1.3: Effects of Apologies where Contributions Decrease after Transgression

<table>
<thead>
<tr>
<th>Contribution Beliefs</th>
<th>Niceness Beliefs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apology Received 2 Rounds Ago</td>
<td>-0.476</td>
</tr>
<tr>
<td>(0.346)</td>
<td>(0.124)</td>
</tr>
<tr>
<td>Private</td>
<td>-1.069***</td>
</tr>
<tr>
<td>(0.362)</td>
<td>(0.134)</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.329</td>
</tr>
<tr>
<td>(1.293)</td>
<td>(0.506)</td>
</tr>
</tbody>
</table>

Past Two Rounds’ Controls ✓ ✓
Session, Round Fixed Effects ✓ ✓
Observations 660 660

* 0.10 ** 0.05 *** 0.01. Tobit regressions with individual-partner-round level observations. Standard errors in parentheses clustered at the group level. Sample is restricted to cases where a partner did not contribute the top in the group two rounds ago, and decreased contributions one round ago. Controls: Apologies Received, Contribution of partner, Contribution of others besides partner, contribution rank, Partner’s contribution rank, Coordination, Belief — all for past two rounds. Apologies are restricted to those made at non-highest rank.

Table A.1.4: Effects of Apologies where Contributions Decrease after Transgression in Control

<table>
<thead>
<tr>
<th>Contribution Beliefs</th>
<th>Niceness Beliefs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apology received 2 rounds ago</td>
<td>-0.059</td>
</tr>
<tr>
<td>[0.477]</td>
<td>[0.470]</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.271</td>
</tr>
<tr>
<td>[0.177]</td>
<td>[0.000]</td>
</tr>
</tbody>
</table>

Past Two Rounds’ Controls ✓ ✓
Session, Round Fixed Effects ✓ ✓
Observations 660 660

* 0.10 ** 0.05 *** 0.01. Tobit regressions with individual-partner-round level observations. Standard errors in parentheses clustered at the group level. Sample is restricted to cases where a partner did not contribute the top in the group two rounds ago, and decreased contributions one round ago. Estimates are based on the means of 2000 sets of apology simulations in the control group. Controls: Apologies Received, Contribution of partner, Contribution of others besides partner, contribution rank, Partner’s contribution rank, Coordination, Belief — all for past two rounds. Apologies are restricted to those made at non-highest rank.
Table A.1.5: Sending “Apologies” in Control: Contributions, Robustness

<table>
<thead>
<tr>
<th>Dep Var: Sending “Apologies” in Control</th>
<th>Contribution</th>
<th>Simulation based on:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Public</td>
</tr>
<tr>
<td>No. of Apologies Sent</td>
<td>-0.033</td>
<td>-0.280</td>
</tr>
<tr>
<td></td>
<td>[0.549]</td>
<td>[0.426]</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.554*</td>
<td>-2.470*</td>
</tr>
<tr>
<td></td>
<td>[0.075]</td>
<td>[0.073]</td>
</tr>
<tr>
<td>Observations</td>
<td>1140</td>
<td>1140</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Tobit regressions with individual-round level observations. Standard errors clustered at the group level. P-values in square parentheses. Estimates are based on the means of 2000 sets of apology simulations in the control group. Controls: Past contribution, Avg of others’ past contributions, Past contribution rank, Past Coordination. Apologies are restricted to those made at non-highest rank.

Table A.1.6: Effects of Apologies in Control: Beliefs, Robustness

<table>
<thead>
<tr>
<th>Dep Var: Effects of Apologies in Control</th>
<th>Contribution Belief</th>
<th>Niceness Belief</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Public</td>
<td>Private</td>
</tr>
<tr>
<td>Apology Sent</td>
<td>-0.101</td>
<td>-0.321</td>
</tr>
<tr>
<td></td>
<td>[0.469]</td>
<td>[0.299]</td>
</tr>
<tr>
<td>Apology Received</td>
<td>0.089</td>
<td>-0.293</td>
</tr>
<tr>
<td></td>
<td>[0.505]</td>
<td>[0.310]</td>
</tr>
<tr>
<td>No. of 3rd Party Apologies Received by Partner</td>
<td>-0.004</td>
<td>-0.240</td>
</tr>
<tr>
<td></td>
<td>[0.563]</td>
<td>[0.294]</td>
</tr>
<tr>
<td>Constant</td>
<td>-3.642**</td>
<td>-3.354**</td>
</tr>
<tr>
<td></td>
<td>[0.017]</td>
<td>[0.015]</td>
</tr>
<tr>
<td>Past Round Controls</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Session, Round Fixed Effects</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>3420</td>
<td>3420</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Tobit regressions with individual-partner-round level observations. Standard errors clustered at the group level. P-values in square parentheses. Estimates are based on the means of 2000 sets of apology simulations in the control group. Controls: Past contribution of partner, Past contribution of others besides partner, Past contribution rank, Partner’s past contribution rank, Past coordination, Past belief. Apologies are restricted to those made at non-highest rank.
### Table A.1.7: Effects of Partner’s Usage of Apologies on Beliefs in Control, Robustness

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Contribution Belief</th>
<th>Niceness Belief</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Public</td>
<td>Private</td>
</tr>
<tr>
<td>Partner’s Past Sincerity</td>
<td>0.692</td>
<td>0.541</td>
</tr>
<tr>
<td></td>
<td>[0.165]</td>
<td>[0.278]</td>
</tr>
<tr>
<td>Partner’s Past Admission of Errors</td>
<td>-0.181</td>
<td>-0.310</td>
</tr>
<tr>
<td></td>
<td>[0.478]</td>
<td>[0.440]</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.292</td>
<td>-2.119</td>
</tr>
<tr>
<td></td>
<td>[0.114]</td>
<td>[0.170]</td>
</tr>
<tr>
<td>Past Round Controls and Apologies</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Session, Round Fixed Effects</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>2432</td>
<td>2432</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Tobit regressions with individual-partner-round level observations. Subset of observations where transgressions and apologies have been observed in the past. Standard errors clustered at the group level. P-values in square parentheses. Estimates are based on the means of 2000 sets of apology simulations in the control group. Past Admission of Errors is the proportion of rounds in the past where a partner apologises when contributing strictly less. Past Sincerity is the proportion of rounds in the past where a partner strictly increases contributions after apologising.

Controls: Past contribution of partner, Past contribution of others besides partner, Past contribution rank, Partner’s past contribution rank, Past coordination, Past belief, Standard apologies received and sent (× treatments).

### Table A.1.8: Effects of Group Usage of Apologies in Control, Robustness

<table>
<thead>
<tr>
<th>Dep Var:</th>
<th>Contribution Belief</th>
<th>Niceness Belief</th>
<th>Contributions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Public</td>
<td>Private</td>
<td>Public</td>
</tr>
<tr>
<td>Group Norm of</td>
<td>1.228</td>
<td>0.988</td>
<td>0.190</td>
</tr>
<tr>
<td>Sincerity</td>
<td>[0.108]</td>
<td>[0.353]</td>
<td>[0.470]</td>
</tr>
<tr>
<td>Group Norm of</td>
<td>2.903</td>
<td>0.050</td>
<td>0.755</td>
</tr>
<tr>
<td>Admitting Error</td>
<td>[0.274]</td>
<td>[0.472]</td>
<td>[0.251]</td>
</tr>
<tr>
<td>Constant</td>
<td>-5.375**</td>
<td>-3.787*</td>
<td>0.761</td>
</tr>
<tr>
<td></td>
<td>[0.016]</td>
<td>[0.054]</td>
<td>[0.212]</td>
</tr>
<tr>
<td>Past Round Controls and Apologies</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Session, Round Fixed Effects</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>3420</td>
<td>3420</td>
<td>3420</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Columns 1 - 4: Tobit regressions with individual-partner-round level observations. Column 3: Tobit regression with individual-round level observations. Standard errors clustered at the group level. P-values in square parentheses. Estimates are based on the means of 2000 sets of apology simulations in the control group. Group norm of admitting error is measured as the proportion of times in the group where an apology is made in cases where a partner has a strictly higher contribution. Group norm of sincerity is measured as the proportion of times in the group where an apology is followed by a strict increase in contributions.

Controls (Columns 1 - 4): Past contribution of partner, Past contribution of others besides partner, Past contribution rank, Partner’s past contribution rank, Past coordination, Past belief, Standard apologies received and sent. Controls (Column 5 - 6): Past contribution, Avg of others’ past contributions, Past contribution rank, Past Coordination, Standard apologies sent.
### Table A.1.9: Personality Traits and Tendency to Apologise

<table>
<thead>
<tr>
<th>Dep Var: Made Apology</th>
<th>Public</th>
<th>Private</th>
</tr>
</thead>
</table>
| First Round Contribution | 0.010  
(0.025) | 0.025  
(0.030) |
| Guilt (NBE)           | -0.040  
(0.139) | 0.216  
(0.136) |
| Guilt (Repair)        | 0.056  
(0.130) | 0.036  
(0.149) |
| Shame (NSE)           | 0.287*  
(0.172) | 0.102  
(0.107) |
| Shame (Withdraw)      | -0.137  
(0.095) | -0.089  
(0.119) |
| Constant              | -1.933***  
(0.618) | -3.215***  
(0.852) |
|                       | -2.391***  
(0.548) | -3.960***  
(0.892) |
| Controls              | ✓ ✓ ✓ ✓ |
| Session, Round Fixed Effects | ✓ ✓ ✓ ✓ |
| Observations          | 3648  | 3648 |

* 0.10 ** 0.05 *** 0.01. Logit regressions with individual-partner-round level observations. Standard errors in parentheses clustered at the group level. Controls: Contribution difference with other, Contribution Rank and their interaction. Guilt and Shame measures are derived from the GASP scale.

### Table A.1.10: Personality Traits and Tendency to Sincerely Apologise

<table>
<thead>
<tr>
<th>Dep Var: Made Sincere Apology</th>
<th>Public</th>
<th>Private</th>
</tr>
</thead>
</table>
| First Round Contribution      | 0.023  
(0.024) | 0.061**  
(0.029) |
| Guilt (NBE)                   | 0.022  
(0.173) | 0.162  
(0.165) |
| Guilt (Repair)                | 0.052  
(0.177) | 0.176  
(0.175) |
| Shame (NSE)                   | 0.352  
(0.230) | 0.156  
(0.125) |
| Shame (Withdraw)              | -0.259***  
(0.083) | 0.020  
(0.136) |
| Constant                      | -5.046***  
(0.940) | -6.568***  
(1.100) |
|                               | -3.839***  
(0.635) | -6.278***  
(1.060) |
| Controls                      | ✓ ✓ ✓ ✓ |
| Session, Round Fixed Effects  | ✓ ✓ ✓ ✓ |
| Observations                  | 3420  | 3600 |

* 0.10 ** 0.05 *** 0.01. Logit regressions with individual-partner-round level observations. Standard errors in parentheses clustered at the group level. Controls: Contribution difference with other, Contribution Rank and their interaction. Guilt and Shame measures are derived from the GASP scale.
Table A.1.11: Importance of Apology over Treatments

<table>
<thead>
<tr>
<th></th>
<th>Dep Var: Apology Importance</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public</td>
<td></td>
<td>4.337***</td>
<td>3.927***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.382)</td>
<td>(0.413)</td>
</tr>
<tr>
<td>Private</td>
<td></td>
<td>3.562***</td>
<td>3.313***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.370)</td>
<td>(0.352)</td>
</tr>
<tr>
<td>Average Group Contribution</td>
<td></td>
<td>0.093**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.044)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td>3.205***</td>
<td>2.256***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.376)</td>
<td>(0.611)</td>
</tr>
<tr>
<td>Session Fixed Effects</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td>226</td>
<td>226</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. OLS regressions with individual level observations. Standard errors in parentheses clustered at the group level. Apology importance is measured on a Likert scale from 0 to 10 with reference to the questions “Would you prefer keeping the opportunity to apologise in the game than having it removed” in the apology treatments and “Would you prefer to have the opportunity to apologise in the game after observing others’ contributions?” in the control treatments. Excludes 6 participants who filled in the wrong survey.
A.2 Experiment Script

A.2.1 Instructions (General Guidelines)

Instructions

General Guidelines

Welcome to the experiment. Please read the instructions below carefully.

Communication between participants is not allowed. Also, please refrain from using any communication devices. If you have any questions at any time, please raise your hand and an experimenter will come over to see you.

Your decisions in the experiment will be anonymous, and your anonymity will be strictly preserved. Participants will interact with each other using only their assigned IDs. For example, you may learn that "Z" contributed $n$ coins; but you will not be told the real name of "Z".

The main experiment will consist of 2 Sections: you will receive instructions right before they commence. In each Section you will have the opportunity to earn "Coins" which can be converted into Pounds at the rate of 100 Coins: £4. At the end of these 2 Sections, you will be asked to fill in a questionnaire which will take approximately 10 minutes. You will receive an additional fee of £2 for completion of the questionnaire. Your final earnings will be the sum of your earnings in the 2 Sections as well as the show up (£3) and completion (£2) fees. You will be paid privately and confidentially.
A.2.2 Instructions (Control)

Instructions: Section 1

We will now begin Section 1 of the experiment. Please read the instructions below carefully.

Introduction

In this Section, you will have to make decisions over 16 Periods. At the beginning of the Section, the computer will randomly assign you to a group of four participants. Also, the computer will randomly select an ID for you, such as “A” or “B” and participants will keep the same ID throughout. This means that your (and others’) IDs and the members of your group will remain the same throughout.

Each period consists of two stages. In Stage 1, you will have to decide how many coins to contribute to a group project and estimate others’ contributions. In Stage 2, you will learn how much the other members of your group contributed and see your earnings for the period.

Stage 1 of a Period

At the beginning of Stage 1, each group member will receive an endowment of 20 Coins. You have to choose how many Coins you want to contribute to a group project, the rest of your Coin endowment will be kept for yourself. Each other group member will make a similar decision simultaneously, that is nobody will be informed about others’ decisions before making their own decision.

Total Group Income from the group project is equal to $2 \times \text{the sum of contributions}$ by all four group members and is evenly shared amongst them (i.e. each group member receives $0.5 \times \text{sum of contributions}$).

Hence, your earnings from this decision task are the sum of two parts:

\[
\begin{align*}
\text{Personal Income} & = (20 - \text{own contribution}) \\
\text{Share of Group Income} & = \frac{2 \times (\text{own contribution} + \text{others contributions})}{4}
\end{align*}
\]

(The number of Coins kept for yourself) \hspace{2cm} (Your quarter of Total Group Income)

Furthermore, you will be asked to estimate the contributions to the project of the other three members of your group during the period. You will be paid for the accuracy of your estimates:

- For each estimate which exactly equals the correct result, you earn 1.5 Coins.
• For each estimate which deviates by one from the correct result, you earn 1 Coin.
• For each estimate which deviates by two from the correct result, you earn 0.5 Coins.
• Estimates which deviate by three or more from the correct result will earn you nothing.

Lastly, you will be asked about your opinion on how much each of the other 3 members cares about others in your group.

Your Stage 1 earnings in the period will thus be the earnings from the decision task + any earnings from the accuracy of your estimates.

**Stage 2 of a Period**

In Stage 2, everyone will observe the contributions of each group member to the project.

In addition, you will also see a summary of your earnings from Stage 1 of the period.

**Payment**

4 periods in Section 1 will be randomly chosen for your payment. Your earnings (in Coins) for Section 1 will be the sum of your earnings from these randomly chosen periods. A payment summary will be shown at the end of the 2 Sections.

If there is anything you are unclear with in the instructions, please raise your hand and we will answer your query privately.

If you have understood the instructions, click next to move on to the practice round.

Note that you will be able to refer to this set of instructions at any time during Section 1 of the experiment.
A.2.3 Screenshots (Practice Round)

Earnings Explanation

A.2.4 Screenshots (Control)

Screen 1
Section 1 Period 1
Contributions

The contributions by your group members to the group project are as follows:

<table>
<thead>
<tr>
<th>Player ID</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>D (Yuu)</td>
<td>0 Coin(s)</td>
</tr>
<tr>
<td>A</td>
<td>0 Coin(s)</td>
</tr>
<tr>
<td>B</td>
<td>0 Coin(s)</td>
</tr>
<tr>
<td>C</td>
<td>0 Coin(s)</td>
</tr>
</tbody>
</table>

Screen 2

Section 1 Period 1
Period Earnings

You earned 20.0 Coin(s) from the decision task and 4.5 Coin(s) from your estimation.
Your total earnings in this period are 24.5 Coin(s).

Screen 3
A.2.5 Instructions (Public Treatment)

Instructions: Section 1

We will now begin Section 1 of the experiment. Please read the instructions below carefully.

Introduction

In this Section, you will have to make decisions over 16 Periods. At the beginning of the Section, the computer will randomly assign you to a group of four participants. Also, the computer will randomly select an ID for you, such as “A” or “B” and participants will keep the same ID throughout. This means that your (and others’) IDs and the members of your group will remain the same throughout.

Each period consists of three stages. In Stage 1, you will have to decide how many coins to contribute to a group project and estimate others’ contributions. In Stage 2, you will learn how much the other members of your group contributed and decide whether to send an apology message to group members. In Stage 3, others’ messages will be revealed and your earnings for the period shown.

Stage 1 of a Period

At the beginning of Stage 1, each group member will receive an endowment of 20 Coins. You have to choose how many Coins you want to contribute to a group project, the rest of your Coin endowment will be kept for yourself. Each other group member will make a similar decision simultaneously, that is nobody will be informed about others’ decisions before making their own decision.

Total Group Income from the group project is equal to $2 \times \text{the sum of contributions}$ by all four group members and is evenly shared amongst them (i.e. each group member receives $0.5 \times \text{sum of contributions}$).

Hence, your earnings from this decision task are the sum of two parts:

\[
\begin{align*}
\text{Personal Income} & \quad + \quad \frac{2 \times (\text{own contribution} + \text{others contributions})}{4} \\
(\text{The number of Coins kept for yourself}) & \quad \text{Share of Group Income} \quad (\text{Your quarter of Total Group Income})
\end{align*}
\]

Furthermore, you will be asked to estimate the contributions to the project of the other three members of your group during the period. You will be paid for the accuracy of your estimates:
• For each estimate which exactly equals the correct result, you earn **1.5 Coins**.
• For each estimate which deviates by one from the correct result, you earn **1 Coin**.
• For each estimate which deviates by two from the correct result, you earn **0.5 Coins**.
• Estimates which deviate by three or more from the correct result will earn you **nothing**.

Lastly, you will be asked about your opinion on how much each of the other 3 members cares about others in your group.

Your Stage 1 earnings in the period will thus be the earnings from the decision task + any earnings from the accuracy of your estimates.

**Stage 2 of a Period**
In Stage 2, everyone will observe the contributions of each group member to the project.

You will then be asked whether you wish to send a message which reads “I am sorry”. This message will be shown to all members of your group.

**Stage 3 of a Period**
In stage 3, you will then get to see messages sent by all members in your group (if any).

In addition, you will also see a summary of your earnings from Stage 1 of the period.

**Payment**
4 periods in Section 1 will be randomly chosen for your payment. Your earnings (in Coins) for Section 1 will be the sum of your earnings from these randomly chosen periods. A payment summary will be shown at the end of the 2 Sections.

If there is anything you are unclear with in the instructions, please raise your hand and we will answer your query privately.

If you have understood the instructions, click next to move on to the practice round.

Note that you will be able to refer to this set of instructions at any time during Section 1 of the experiment.
A.2.6 Screenshots (Public Treatment)

Section 1 Period 1

You have 20 Coin(s).
Please decide how much of this to contribute to the group project:

[Input field for contribution amount]

What are your estimates of each other group member's contribution in this period?
Also, on a scale from 1 to 7, how much do you think each group member cares about others?

<table>
<thead>
<tr>
<th>Player ID</th>
<th>Your estimates</th>
<th>Care about others</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Submit

Screen 1

Section 1 Period 1
Contributions

The contributions by your group members to the group project are as follows:

<table>
<thead>
<tr>
<th>Player ID</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>D (You)</td>
<td>0 Coin(s)</td>
</tr>
<tr>
<td>A</td>
<td>0 Coin(s)</td>
</tr>
<tr>
<td>B</td>
<td>0 Coin(s)</td>
</tr>
<tr>
<td>C</td>
<td>0 Coin(s)</td>
</tr>
</tbody>
</table>

Would you like to send a message which reads "I am sorry" to your group?

[Message field]

Submit

Screen 2
Section 1 Period 1
Messages

The messages sent by your group members are as follows:

<table>
<thead>
<tr>
<th>Player ID</th>
<th>Contribution</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>D (You)</td>
<td>0 Coin(s)</td>
<td>I am sorry.</td>
</tr>
<tr>
<td>A</td>
<td>0 Coin(s)</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0 Coin(s)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0 Coin(s)</td>
<td></td>
</tr>
</tbody>
</table>

Screen 3

Section 1 Period 1
Period Earnings

You earned 20.0 Coin(s) from the decision task and 4.5 Coin(s) from your estimation.
Your total earnings in this period are 24.5 Coin(s).

Screen 4
Instructions: Section 1

We will now begin Section 1 of the experiment. Please read the instructions below carefully.

Introduction

In this Section, you will have to make decisions over 16 Periods. At the beginning of the Section, the computer will randomly assign you to a group of four participants. Also, the computer will randomly select an ID for you, such as “A” or “B” and participants will keep the same ID throughout. This means that your (and others’) IDs and the members of your group will remain the same throughout.

Each period consists of three stages. In Stage 1, you will have to decide how many coins to contribute to a group project and estimate others’ contributions. In Stage 2, you will learn how much the other members of your group contributed and decide whether to send an apology message to group members. In Stage 3, others’ messages will be revealed and your earnings for the period shown.

Stage 1 of a Period

At the beginning of Stage 1, each group member will receive an endowment of 20 Coins. You have to choose how many Coins you want to contribute to a group project, the rest of your Coin endowment will be kept for yourself. Each other group member will make a similar decision simultaneously, that is nobody will be informed about others’ decisions before making their own decision.

Total Group Income from the group project is equal to $2 \times \text{the sum of contributions}$ by all four group members and is evenly shared amongst them (i.e. each group member receives $0.5 \times \text{sum of contributions}$).

Hence, your earnings from this decision task are the sum of two parts:

\[
\begin{align*}
\text{Personal Income} & \quad + \quad \frac{2 \times (\text{own contribution} + \text{others contributions})}{4} \\
(\text{The number of Coins kept for yourself}) & \quad \text{Share of Group Income} \quad (\text{Your quarter of Total Group Income})
\end{align*}
\]

Furthermore, you will be asked to estimate the contributions to the project of the other three members of your group during the period. You will be paid for the accuracy of your estimates:
• For each estimate which exactly equals the correct result, you earn 1.5 Coins.
• For each estimate which deviates by one from the correct result, you earn 1 Coin.
• For each estimate which deviates by two from the correct result, you earn 0.5 Coins.
• Estimates which deviate by three or more from the correct result will earn you nothing.

Lastly, you will be asked about your opinion on how much each of the other 3 members cares about others in your group.

Your Stage 1 earnings in the period will thus be the earnings from the decision task + any earnings from the accuracy of your estimates.

Stage 2 of a Period
In Stage 2, everyone will observe the contributions of each group member to the project.

For each other group member, you will then be asked whether you wish to send a message which reads "I am sorry". Messages will only be seen by the target recipient.

Stage 3 of a Period
In stage 3, you will then get to see messages sent to you by others in your group (if any).

In addition, you will also see a summary of your earnings from Stage 1 of the period.

Payment
4 periods in Section 1 will be randomly chosen for your payment. Your earnings (in Coins) for Section 1 will be the sum of your earnings from these randomly chosen periods. A payment summary will be shown at the end of the 2 Sections.

If there is anything you are unclear with in the instructions, please raise your hand and we will answer your query privately.

If you have understood the instructions, click next to move on to the practice round.

Note that you will be able to refer to this set of instructions at any time during Section 1 of the experiment.
A.2.8 Screenshots (Private Treatment)

**Screen 1**

Section 1 Period 1

You have 20 Coin(s).
Please decide how much of this to contribute to the group project:

What are your estimates of each other group member’s contribution in this period?
Also, on a scale from 1 to 7, how much do you think each group member cares about others?

<table>
<thead>
<tr>
<th>Player ID</th>
<th>Your estimates</th>
<th>Cares about others</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Coin 10</td>
<td>Not at all</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>C</td>
<td>Coin 0</td>
<td>Not at all</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>D</td>
<td>Coin 0</td>
<td>Not at all</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 2 3 4 5 6 7</td>
</tr>
</tbody>
</table>

Submit

**Screen 2**

Section 1 Period 1
Contributions

The contributions by your group members to the group project are as follows:

<table>
<thead>
<tr>
<th>Player ID</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>0 Coin(s)</td>
</tr>
<tr>
<td>A</td>
<td>0 Coin(s)</td>
</tr>
<tr>
<td>C</td>
<td>0 Coin(s)</td>
</tr>
<tr>
<td>D</td>
<td>0 Coin(s)</td>
</tr>
</tbody>
</table>

Would you like to send a message which reads "I am sorry" to Player A?

Would you like to send a message which reads "I am sorry" to Player C?

Would you like to send a message which reads "I am sorry" to Player D?

Submit
Section 1 Period 1
Messages

Your message(s) have been delivered to their recipients.

The messages sent to you by other group members are as follows:

<table>
<thead>
<tr>
<th>Player ID</th>
<th>Contribution</th>
<th>Message to you</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0 Coin(s)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0 Coin(s)</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>0 Coin(s)</td>
<td></td>
</tr>
</tbody>
</table>

Next

Screen 3

Section 1 Period 1
Period Earnings

You earned 20.0 Coin(s) from the decision task and 4.5 Coin(s) from your estimation.

Your total earnings in this period are 24.5 Coin(s).

Next

Screen 4
Section 1 of the experiment has ended. We will now begin Section 2 of the experiment. Please read the instructions below carefully.

In this Section, **the computer will again randomly assign you to a new group of four participants** and you will play the same game (same endowment and formula for earnings calculation) as before **for only 1 Period**. However, the way contributions are decided will be different from before.

For this period only, you will be asked to give two type of contribution decisions. Firstly, you will be asked to choose how many Coins you want to contribute to the group project as before. This is your **unconditional** contribution.

Secondly, you will be asked to fill in a contribution table as below where you indicate how many Coins you wish to contribute to the project for each possible average contribution of other group members (rounded to the nearest integer). This is your **conditional** contribution.

<table>
<thead>
<tr>
<th>Coin(s)</th>
<th>Coin(s)</th>
<th>Coin(s)</th>
<th>Coin(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>7</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>

After all participants have made their unconditional and conditional contributions, 1 group member will be randomly selected from your group. The three other group members’ contribution will be their **unconditional contribution**; The randomly selected group member’s contribution will be his/her **conditional contribution** given the average of the other three’s (unconditional) contributions as in his/her contribution table.

Once everyone’s contributions have been determined as above, the earnings are calculated similarly to before. You obviously do not know whether the random mechanism will select you when you make your unconditional contribution and when
you fill in the contribution table. **You will therefore have to think carefully about both types of decisions because both can become relevant for you.**

Your earnings from this period will be added to your earnings from Section 1 to become your total earnings.

You can click the button below to see an example to help you understand the rules of this section. If there is anything you are unclear with in the instructions, please raise your hand and we will answer your query privately. If you have understood the instructions, click next to move on.

**Example**

Assume that there are 4 players W,X,Y,Z. Player Z is randomly selected by the computer. This means that Player Z’s relevant decision will be the conditional contribution in his/her contribution table. The unconditional contribution is the relevant decision for Player W, X and Y.

Assume that Players W,X and Y made unconditional contributions of 0, 2, and 4 Coins. The average contribution of these three group members, therefore, is 2 Coins.

If Player Z indicated in his/her contribution table that he/she would contribute 1 token if the others contribute 2 Coins on average, then the total contribution to the project is given by 0+2+4+1=7 Coins. All group members, therefore, earn (2×7)/4=3.5 Coins from the project plus the number of Coins kept for themselves.

If, instead, Player Z indicated in his/her contribution table that he/she would contribute 19 Coins if the others contribute two Coins on average, then the total contribution of the group to the project is given by 0+2+4+19=25. All group members therefore earn (2×25)/4=12.5 points from the project plus the number of Coins kept for themselves.
### A.2.10 Screenshots (Type Elicitation)

**Screen 1**

![Screen 1](image1)

**Screen 2**

![Screen 2](image2)

<table>
<thead>
<tr>
<th>6</th>
<th>Coin(s)</th>
<th>7</th>
<th>Coin(s)</th>
<th>14</th>
<th>Coin(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Coin(s)</td>
<td>8</td>
<td>Coin(s)</td>
<td>15</td>
<td>Coin(s)</td>
</tr>
<tr>
<td>2</td>
<td>Coin(s)</td>
<td>9</td>
<td>Coin(s)</td>
<td>16</td>
<td>Coin(s)</td>
</tr>
<tr>
<td>3</td>
<td>Coin(s)</td>
<td>10</td>
<td>Coin(s)</td>
<td>17</td>
<td>Coin(s)</td>
</tr>
<tr>
<td>4</td>
<td>Coin(s)</td>
<td>11</td>
<td>Coin(s)</td>
<td>18</td>
<td>Coin(s)</td>
</tr>
<tr>
<td>5</td>
<td>Coin(s)</td>
<td>12</td>
<td>Coin(s)</td>
<td>19</td>
<td>Coin(s)</td>
</tr>
<tr>
<td>6</td>
<td>Coin(s)</td>
<td>13</td>
<td>Coin(s)</td>
<td>20</td>
<td>Coin(s)</td>
</tr>
</tbody>
</table>
A.2.11 Post-Experiment Survey Questions

Demographic questions
What is your age? (If you would prefer not to answer, please leave it blank.)
What is your year of study?
[1st Year, 2nd Year, 3rd Year, 4th Year, 5th Year, Postgraduate, Other]
What is your nationality?
What is your course of study?
What is your gender?
[Male, Female, I’d prefer not to answer, Other (Please describe if you wish)]

Miscellaneous questions
How would you rate your happiness at the moment? [1: Completely Sad - 7: Completely Happy]

Game behaviour*
Was there anything unclear about the instructions? If so, please state below.
How did you decide how much to contribute?
How did you decide whether to apologise? Did you think you would be forgiven when giving an apology?†
Suppose in one of the rounds, you gave less than the other members. How do you think an apology would affect other group members’ expectations of how much you would give in the next round?†
If you could add to the message of apology, what would you write?†
Did observing an apology from another player change your decision on contribution in the next round? If so, how?†
How did you feel when receiving an apology from another player? (if any)†
On a scale from 0 to 10, would you prefer keeping the opportunity to apologise in the game than having it removed? [0: Not at all - 10: Yes, a lot] †
On a scale from 0 to 10, would you prefer to have the opportunity to apologise in the game after observing others’ contributions? [0: Not at all - 10: Yes, a lot] ‡
Did you find anything about the experiment unfair? If so, please describe.

Preference questions**
Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people? [0: Need to be very careful - 10: Most people can be trusted]
Please tell me, in general, how willing or unwilling you are to take risks.
[0: Completely unwilling to take risks - 10: Very willing to take risks]
When someone does me a favor I am willing to return it.
[0: Does not describe me at all - 10: Describes me perfectly]
If I am treated very unjustly, I will take revenge at the first occasion, even if there is a cost to do so.
[0: Does not describe me at all - 10: Describes me perfectly]
How willing are you to give to good causes without expecting anything in return?
[0: Completely unwilling to do so - 10: Very willing to do so]
I assume that people have only the best intentions.
[0: Does not describe me at all - 10: Describes me perfectly]
Guilt and shame aversion (Cohen et al., 2011)**

Over the next few pages, you will read about situations that people are likely to encounter in day-to-day life, followed by common reactions to those situations. As you read each scenario, try to imagine yourself in that situation. Then indicate the likelihood that you would react in the way described.

[1: Very Unlikely - 7: Very Likely ]

**Guilt-Negative-Behavior-Evaluation (NBE)**

After realizing you have received too much change at a store, you decide to keep it because the sales clerk doesn’t notice. What is the likelihood that you would feel uncomfortable about keeping the money?

You secretly commit a felony. What is the likelihood that you would feel remorse about breaking the law?

At a coworker’s housewarming party, you spill red wine on their new cream-colored carpet. You cover the stain with a chair so that nobody notices your mess. What is the likelihood that you would feel that the way you acted was pathetic?

You lie to people but they never find out about it. What is the likelihood that you would feel terrible about the lies you told?

**Guilt-Repair**

You are privately informed that you are the only one in your group that did not make the honor society because you skipped too many days of school. What is the likelihood that this would lead you to become more responsible about attending school?

You reveal a friend’s secret, though your friend never finds out. What is the likelihood that your failure to keep the secret would lead you to exert extra effort to keep secrets in the future?

You strongly defend a point of view in a discussion, and though nobody was aware of it, you realize that you were wrong. What is the likelihood that this would make you think more carefully before you speak?

While discussing a heated subject with friends, you suddenly realize you are shouting though nobody seems to notice. What is the likelihood that you would try to act more considerately toward your friends?

**Shame-Negative-Self-Evaluation (NSE)**

You rip an article out of a journal in the library and take it with you. Your teacher discovers what you did and tells the librarian and your entire class. What is the likelihood that this would make you feel like a bad person?

You give a bad presentation at work. Afterwards your boss tells your coworkers it was your fault that your company lost the contract. What is the likelihood that you would feel incompetent?

You successfully exaggerate your damages in a lawsuit. Months later, your lies are discovered and you are charged with perjury. What is the likelihood that you would think you are a despicable human being?

You make a mistake at work and find out a coworker is blamed for the error. Later, your coworker confronts you about your mistake. What is the likelihood that you would feel like a coward?
Guilt and shame aversion (Continued)

Shame-Withdraw

After making a big mistake on an important project at work in which people were depending on you, your boss criticizes you in front of your coworkers. What is the likelihood that you would feign sickness and leave work?

A friend tells you that you boast a great deal. What is the likelihood that you would stop spending time with that friend?

Your home is very messy and unexpected guests knock on your door and invite themselves in. What is the likelihood that you would avoid the guests until they leave?

You take office supplies home for personal use and are caught by your boss. What is the likelihood that this would lead you to quit your job?

*6 participants filled in the incorrect ID and thus received the wrong open-ended portion of the post-experiment survey. The rest of the survey is independent of participants’ treatment. In particular, 2 in the control received questions meant for the apology treatment while 4 in the apology treatment received questions meant for the control. We exclude responses on these questions from any analysis.

**Order of questions was randomised within section.

† Only included in survey for apology treatments.

‡ Only included in survey for control treatment.
Appendix B (for Chapter 2)

B.1 Appendix Figures and Tables

Table B.1.1: Correlation between ingroup bias and survey-measured feelings of belonging

<table>
<thead>
<tr>
<th>Dep var: Ingroup Bias</th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feelings of belonging to ingroup</td>
<td>2.987***</td>
</tr>
<tr>
<td></td>
<td>(0.473)</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.338</td>
</tr>
<tr>
<td></td>
<td>(3.387)</td>
</tr>
<tr>
<td>Strata, Session and Order Fixed effects</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>1437</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Ordinary least squares regressions with individual-question level observations. Standard error in parentheses clustered at the individual level.

Table B.1.2: Effects of multiple salient dimensions on ingroup correlation beliefs

<table>
<thead>
<tr>
<th>Dep var: Ingroup correlation beliefs</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 salient dimensions</td>
<td>-3.099</td>
<td>-4.043</td>
</tr>
<tr>
<td></td>
<td>(2.334)</td>
<td>(2.845)</td>
</tr>
<tr>
<td>Government dimension</td>
<td>-5.727*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.437)</td>
<td></td>
</tr>
<tr>
<td>2 salient dimensions × Government dimension</td>
<td>1.878</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.759)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>7.069***</td>
<td>9.759**</td>
</tr>
<tr>
<td></td>
<td>(1.786)</td>
<td>(2.325)</td>
</tr>
<tr>
<td>Strata, Session and Order Fixed effects</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>1232</td>
<td>1232</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Ordinary least squares regressions with individual-question level observations. Out of range beliefs are treated as missing. Standard error in parentheses clustered at the individual level.
Table B.1.3: Sample Characteristics

<table>
<thead>
<tr>
<th></th>
<th>1-salient-dimension mean (sd)</th>
<th>2-salient-dimensions mean (sd)</th>
<th>Difference (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>35.730 (13.177)</td>
<td>35.624 (12.993)</td>
<td>0.106 (0.900)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.493 (0.500)</td>
<td>0.475 (0.500)</td>
<td>0.018 (0.577)</td>
</tr>
<tr>
<td>Male</td>
<td>0.489 (0.500)</td>
<td>0.498 (0.501)</td>
<td>-0.009 (0.775)</td>
</tr>
<tr>
<td>Other or preferred not to say</td>
<td>0.019 (0.135)</td>
<td>0.027 (0.163)</td>
<td>-0.009 (0.366)</td>
</tr>
</tbody>
</table>

*Means here reflect proportions within the corresponding population.*

<table>
<thead>
<tr>
<th><strong>Highest education level completed</strong></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than High School Diploma</td>
<td>0.019 (0.135)</td>
<td>0.008 (0.091)</td>
<td>0.010 (0.172)</td>
</tr>
<tr>
<td>High School Diploma</td>
<td>0.351 (0.478)</td>
<td>0.326 (0.469)</td>
<td>0.025 (0.415)</td>
</tr>
<tr>
<td>Associate’s Degree (AA/AS)</td>
<td>0.076 (0.266)</td>
<td>0.118 (0.323)</td>
<td>-0.041* (0.030)</td>
</tr>
<tr>
<td>Bachelor’s Degree (BA/BS)</td>
<td>0.386 (0.487)</td>
<td>0.370 (0.483)</td>
<td>0.016 (0.613)</td>
</tr>
<tr>
<td>Master’s Degree (MA/MS/M.Eng)</td>
<td>0.140 (0.348)</td>
<td>0.143 (0.350)</td>
<td>-0.003 (0.906)</td>
</tr>
<tr>
<td>Doctoral Degree (PhD)</td>
<td>0.012 (0.111)</td>
<td>0.017 (0.129)</td>
<td>-0.004 (0.567)</td>
</tr>
<tr>
<td>Professional Degree (e.g., JD, MD)</td>
<td>0.016 (0.128)</td>
<td>0.019 (0.136)</td>
<td>-0.002 (0.777)</td>
</tr>
</tbody>
</table>

*Means here reflect proportions within the corresponding population.*

Continued on next page...
<table>
<thead>
<tr>
<th>Annual personal income</th>
<th>1-salient-dimension mean (sd)</th>
<th>2-salient-dimensions mean (sd)</th>
<th>Difference (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under $10,000</td>
<td>0.270 (0.444)</td>
<td>0.242 (0.429)</td>
<td>0.029 (0.312)</td>
</tr>
<tr>
<td>$10,000 to $19,999</td>
<td>0.124 (0.330)</td>
<td>0.132 (0.339)</td>
<td>-0.009 (0.689)</td>
</tr>
<tr>
<td>$20,000 to $29,999</td>
<td>0.109 (0.312)</td>
<td>0.109 (0.312)</td>
<td>0.000 (0.999)</td>
</tr>
<tr>
<td>$30,000 to $39,999</td>
<td>0.107 (0.310)</td>
<td>0.097 (0.296)</td>
<td>0.011 (0.588)</td>
</tr>
<tr>
<td>$40,000 to $49,999</td>
<td>0.076 (0.266)</td>
<td>0.088 (0.284)</td>
<td>-0.012 (0.501)</td>
</tr>
<tr>
<td>$50,000 to $59,999</td>
<td>0.074 (0.262)</td>
<td>0.065 (0.247)</td>
<td>0.009 (0.580)</td>
</tr>
<tr>
<td>$60,000 to $69,999</td>
<td>0.035 (0.184)</td>
<td>0.053 (0.223)</td>
<td>-0.017 (0.187)</td>
</tr>
<tr>
<td>$70,000 to $79,999</td>
<td>0.054 (0.225)</td>
<td>0.057 (0.232)</td>
<td>-0.003 (0.833)</td>
</tr>
<tr>
<td>$80,000 to $89,999</td>
<td>0.031 (0.173)</td>
<td>0.040 (0.196)</td>
<td>-0.009 (0.452)</td>
</tr>
<tr>
<td>$90,000 to $99,999</td>
<td>0.025 (0.155)</td>
<td>0.021 (0.144)</td>
<td>0.004 (0.699)</td>
</tr>
<tr>
<td>$100,000 to $124,999</td>
<td>0.047 (0.213)</td>
<td>0.034 (0.180)</td>
<td>0.014 (0.278)</td>
</tr>
<tr>
<td>$125,000 to $149,999</td>
<td>0.012 (0.111)</td>
<td>0.017 (0.129)</td>
<td>-0.004 (0.567)</td>
</tr>
<tr>
<td>$150,000 and above</td>
<td>0.035 (0.184)</td>
<td>0.046 (0.210)</td>
<td>-0.011 (0.382)</td>
</tr>
</tbody>
</table>

Means here reflect proportions within the corresponding population.
<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>1-salient-dimension mean (sd)</th>
<th>2-salient-dimensions mean (sd)</th>
<th>Difference (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>0.693 (0.462)</td>
<td>0.662 (0.474)</td>
<td>0.031 (0.304)</td>
</tr>
<tr>
<td>Black or African American</td>
<td>0.072 (0.259)</td>
<td>0.101 (0.301)</td>
<td>-0.029 (0.114)</td>
</tr>
<tr>
<td>Native American or American Indian</td>
<td>0.004 (0.064)</td>
<td>0.000 (0.000)</td>
<td>0.004 (0.158)</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>0.064 (0.245)</td>
<td>0.063 (0.243)</td>
<td>0.001 (0.955)</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>0.126 (0.332)</td>
<td>0.143 (0.350)</td>
<td>-0.017 (0.438)</td>
</tr>
<tr>
<td>Others or Multiracial</td>
<td>0.041 (0.199)</td>
<td>0.032 (0.175)</td>
<td>0.010 (0.421)</td>
</tr>
</tbody>
</table>

*Means here reflect proportions within the corresponding population.*

<table>
<thead>
<tr>
<th>If understood instructions</th>
<th>1-salient-dimension mean (sd)</th>
<th>2-salient-dimensions mean (sd)</th>
<th>Difference (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.988 (0.111)</td>
<td>0.987 (0.112)</td>
<td>0.000 (0.974)</td>
</tr>
<tr>
<td>No</td>
<td>0.012 (0.111)</td>
<td>0.013 (0.112)</td>
<td>-0.000 (0.974)</td>
</tr>
</tbody>
</table>

*Means here reflect proportions within the corresponding population.*

Continued on next page...
### Made allocation decisions

<table>
<thead>
<tr>
<th>Made allocation decisions</th>
<th>1-salient-dimension mean (sd)</th>
<th>2-salient-dimensions mean (sd)</th>
<th>Difference (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Based on my preferences</td>
<td>0.404 (0.491)</td>
<td>0.487 (0.500)</td>
<td>-0.083***</td>
</tr>
<tr>
<td>Based on what I wanted to do</td>
<td>0.223 (0.416)</td>
<td>0.200 (0.400)</td>
<td>0.023</td>
</tr>
<tr>
<td>Based on what I thought the experimenter expected me to do</td>
<td>0.004 (0.064)</td>
<td>0.006 (0.079)</td>
<td>-0.002</td>
</tr>
<tr>
<td>Randomly</td>
<td>0.010 (0.101)</td>
<td>0.011 (0.102)</td>
<td>0.000</td>
</tr>
<tr>
<td>Based on gut instinct</td>
<td>0.177 (0.382)</td>
<td>0.130 (0.337)</td>
<td>0.047*</td>
</tr>
<tr>
<td>Based on what I thought was the socially desired thing to do</td>
<td>0.047 (0.213)</td>
<td>0.034 (0.180)</td>
<td>0.014</td>
</tr>
<tr>
<td>Based on what I thought was fair†</td>
<td>0.097 (0.296)</td>
<td>0.078 (0.268)</td>
<td>0.019</td>
</tr>
<tr>
<td>Other</td>
<td>0.037 (0.189)</td>
<td>0.055 (0.227)</td>
<td>-0.018</td>
</tr>
</tbody>
</table>

*Means here reflect proportions within the corresponding population.*

Sample size (n) 485 476

*0.10 **0.05 ***0.01. Standard deviations in parentheses for first two columns. P-values for t-tests in parentheses for third column.

†This option only appeared in the third session.
Table B.1.4: Marginal effects of two salient dimensions, relative to one salient dimension, on raw ingroup bias in each identity dimension

<table>
<thead>
<tr>
<th>Subgroup:</th>
<th>Religion Dimension</th>
<th>Government Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Marginal Effect on Raw Ingroup Bias of:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two Salient Dimensions</td>
<td>11.336***</td>
<td>-0.521</td>
</tr>
<tr>
<td></td>
<td>(3.269)</td>
<td>(2.938)</td>
</tr>
<tr>
<td>Observations</td>
<td>1437</td>
<td></td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Estimates are from an ordinary least squares regression using individual-question level observations. It regresses raw ingroup bias on the interaction between dummies for the number of salient dimensions and dummies for the dimension of allocation. Includes controls for strata, session and order fixed effects. Standard errors clustered at individual level in parentheses.

Table B.1.5: Marginal effects of two salient dimensions relative to one salient dimension, and ingroup correlation beliefs, on raw ingroup bias in each identity dimension

<table>
<thead>
<tr>
<th>Subgroup:</th>
<th>Religion Dimension</th>
<th>Government Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Marginal Effects on Raw Ingroup Bias of:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 salient dimensions</td>
<td>11.295***</td>
<td>-5.955*</td>
</tr>
<tr>
<td></td>
<td>(3.314)</td>
<td>(3.307)</td>
</tr>
<tr>
<td>Ingroup correlation beliefs</td>
<td>0.184***</td>
<td>-0.015</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>Observations</td>
<td>1232</td>
<td></td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Estimates are from an ordinary least squares regression using individual-question level observations. It regresses raw ingroup bias on the interaction between dummies for the number of salient dimensions, dummies for the dimension of allocation and ingroup correlation beliefs. 215 observations excluded due to out of range beliefs which are treated as missing. Includes controls for strata, session and order fixed effects. Standard errors clustered at individual level in parentheses.

Table B.1.6: Marginal effects of feelings of belonging to ingroup, on raw ingroup bias in each identity dimension

<table>
<thead>
<tr>
<th>Subgroup:</th>
<th>Religion Dimension</th>
<th>Government Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Marginal Effect on Raw Ingroup Bias of:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feelings of belonging to ingroup</td>
<td>2.572***</td>
<td>4.676***</td>
</tr>
<tr>
<td></td>
<td>(0.636)</td>
<td>(0.783)</td>
</tr>
<tr>
<td>Observations</td>
<td>1437</td>
<td></td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Estimates are from an ordinary least squares regression using individual-question level observations. It regresses raw ingroup bias on the interaction between feelings of belonging to ingroup and dummies for the dimension of allocation. Includes controls for strata, session and order fixed effects. Standard errors clustered at individual level in parentheses.
Table B.1.7: Marginal effects of two salient dimensions, relative to one salient dimension, on ingroup bias in each identity dimension

<table>
<thead>
<tr>
<th>Subgroup:</th>
<th>Religion Dimension</th>
<th>Government Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marginal Effect on ingroup bias of:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two Salient Dimensions</td>
<td>7.424***</td>
<td>0.101</td>
</tr>
<tr>
<td></td>
<td>(3.067)</td>
<td>(2.752)</td>
</tr>
</tbody>
</table>

Observations: 1437

* 0.10 ** 0.05 *** 0.01. Estimates are from an ordinary least squares regression using individual-question level observations. It regresses ingroup bias on the interaction between dummies for the number of salient dimensions and dummies for the dimension of allocation. Includes controls for strata, session, order fixed effects as well as whether the participant indicated they understood instructions, whether the participant indicated they made allocation decisions randomly or based on what they thought the experimenter expected them to do, the participant's age, gender, highest education level completed, personal income level and ethnicity. Standard errors clustered at individual level in parentheses.

Table B.1.8: Effects of multiple salient dimensions by identity dimension on survey-measured feelings of belonging

<table>
<thead>
<tr>
<th>Dep Var: Feelings of belonging to ingroup</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government dimension</td>
<td>-0.577***</td>
<td>-0.682***</td>
</tr>
<tr>
<td></td>
<td>(0.123)</td>
<td>(0.215)</td>
</tr>
<tr>
<td>2 salient dimensions × Religion dimension</td>
<td>-0.094</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.229)</td>
<td></td>
</tr>
<tr>
<td>2 salient dimensions × Government dimension</td>
<td>0.063</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.183)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>7.242***</td>
<td>7.320***</td>
</tr>
<tr>
<td></td>
<td>(0.935)</td>
<td>(0.954)</td>
</tr>
<tr>
<td>Controls</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Strata, Session and Order Fixed effects</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>1437</td>
<td>1437</td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Ordinary least squares regressions with individual-question level observations. Includes controls for strata, session, order fixed effects as well as whether the participant indicated they understood instructions, whether the participant indicated they made allocation decisions randomly or based on what they thought the experimenter expected them to do, the participant’s age, gender, highest education level completed, personal income level and ethnicity. Standard error in parentheses clustered at the individual level.
Table B.1.9: Marginal effects of two salient dimensions relative to one salient dimension, and ingroup correlation beliefs, on ingroup bias in each identity dimension

<table>
<thead>
<tr>
<th>Subgroup:</th>
<th>Religion Dimension</th>
<th>Government Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marginal Effects on Ingroup bias of:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 salient dimensions</td>
<td>8.045***</td>
<td>-4.337</td>
</tr>
<tr>
<td>(3.085)</td>
<td>(3.169)</td>
<td></td>
</tr>
<tr>
<td>Ingroup correlation beliefs</td>
<td>0.191***</td>
<td>0.008</td>
</tr>
<tr>
<td>(0.049)</td>
<td>(0.046)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1232</td>
<td></td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Estimates are from an ordinary least squares regression using individual-question level observations. It regresses ingroup bias on the interaction between dummies for the number of salient dimensions, dummies for the dimension of allocation and ingroup correlation beliefs. 215 observations excluded due to out of range beliefs which are treated as missing. Includes controls for strata, session and order fixed effects as well as whether the participant indicated they understood instructions, whether the participant indicated they made allocation decisions randomly or based on what they thought the experimenter expected them to do, the participant’s age, gender, highest education level completed, personal income level and ethnicity. Standard errors clustered at individual level in parentheses.

Table B.1.10: Marginal effects of feelings of belonging to ingroup, on ingroup bias in each identity dimension

<table>
<thead>
<tr>
<th>Subgroup:</th>
<th>Religion Dimension</th>
<th>Government Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marginal Effect on Ingroup bias of:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feelings of belonging to ingroup</td>
<td>2.176***</td>
<td>3.919***</td>
</tr>
<tr>
<td>(0.602)</td>
<td>(0.739)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1437</td>
<td></td>
</tr>
</tbody>
</table>

* 0.10 ** 0.05 *** 0.01. Estimates are from an ordinary least squares regression using individual-question level observations. It regresses ingroup bias on the interaction between feelings of belonging to ingroup and dummies for the dimension of allocation. Includes controls for strata, session and order fixed effects as well as whether the participant indicated they understood instructions, whether the participant indicated they made allocation decisions randomly or based on what they thought the experimenter expected them to do, the participant’s age, gender, highest education level completed, personal income level and ethnicity. Standard errors clustered at individual level in parentheses.
B.2 Experiment Script

Experiment Introduction

Participant Information Sheet

You have been invited to take part in a research study run by researchers at the University of Warwick. Please read the following statements carefully and answer the question below.

Our commitments and privacy policy

- We never deceive participants. For example, if we inform you that another participant is making a choice on which you can then react, this is indeed the case.
- We keep our promises made to participants. For example, if we promise a certain payment, participants will indeed receive it. In the event that we are responsible for a mistake that is to the disadvantage of participants, we will inform and compensate the respective participants.
- We design, conduct and report our research in accordance with recognized scientific standards and ethical principles.

We adhere to the terms of our privacy policy as stated below:
The data in the participants’ database will only be used for the purpose of the study. There is no link between the personal data in the participants’ database and the data collected during a study. The generated anonymous data will be used for analysis. The end product will be publicly available. Your participation in this study is purely voluntary, and you may withdraw your participation or your data at any time without any penalty to you. Please note that the software automatically notes the time you spent on each question and this data will be made available to researchers for analysis.

Please refer to the University of Warwick Research Privacy Notice which is available here: https://warwick.ac.uk/services/do/dataprotection/privacy/notices/researchers-privacy-notice or by contacting the Information and Data Compliance Team at IDR@warwick.ac.uk.

This study has been reviewed and given favourable opinion by the University of Warwick’s Humanities and Social Science Research Ethics Committee (HSSREC).

If there is anything about the study or your participation that is unclear or that you do not understand, if you have questions or wish to report a research-related problem, you may contact us via message on Proloki.

If you would like to make a complaint about the way you have been dealt with during the study or any possible harm you might have suffered please address your complaint to the person below, who is a senior University of Warwick official entirely independent of this study:

Jane Priest (Head of Research Governance)
Research & Impact Services
University House
University of Warwick
CV4 7JW
Email: researchgovernance@warwick.ac.uk
Tel: 024 76 52466

If you wish to raise a complaint on how we have handled your personal data, you can contact our Data Protection Officer, Anjali Bajaj, Information and Data Director who will investigate the matter: DPO@warwick.ac.uk.

If you are not satisfied with our response or believe we are processing your personal data in a way that is not lawful, you can complain to the Information Commissioner’s Office (ICO).

Thank you for taking the time to read this Participant Information Sheet.

☐ I have read the above and consent to take part in this study
☐ I do not wish to participate

Next

Participation Agreement
Welcome to the study.

In this study, you will be asked to complete several survey questions. There will also be some tasks which earn you credits.

You will interact indirectly with others during the tasks. Your interactions will be anonymous and your anonymity will be strictly preserved.

Credits earned will be converted into US dollars at the rate of 100 Credits: 1 Dollar.

On completing the study, you will be paid a completion fee of $0.90 plus your task earnings.

Please note that this study is best run on a desktop computer. If you are currently using a mobile phone, please close the window and open the study link on a desktop computer.
Identity Priming and Elicitation (2 Salient Dimensions)

Firstly, we would like to ask how you stand on the following issues.

"Government intervention is good for society"
- Agree
- Disagree

"The world was created by a divine entity"
- Agree
- Disagree

Identity Elicitation (2 Salient Dimensions)

Based on your answers on the previous page, you belong to the group which:

"Government intervention is good for society"
- AGREES that:
- DISAGREES that:

"The world was created by a divine entity"
- AGREES that:
- DISAGREES that:

Identity Priming (2 Salient Dimensions)
Identity Priming 2 (2 Salient Dimensions)
Identity Priming and Elicitation (Religion Dimension Salient)

Firstly, we would like to ask how you stand on the following issue.

"The world was created by a divine entity"

- Agree
- Disagree

Identity Elicitation (Religion Dimension Salient)

Based on your answer on the previous page, you belong to the group which:

- AGREES that: "The world was created by a divine entity"
- DISAGREES that: "The world was created by a divine entity"

Identity Priming (Religion Dimension Salient)
Identity Priming 2 (Religion Dimension Salient)
Identity Priming and Elicitation (Government Dimension Salient)

Firstly, we would like to ask how you stand on the following issue.

"Government intervention is good for society"

○ Agree
○ Disagree

Next

Identity Elicitation (Government Dimension Salient)

Based on your answer on the previous page,

You belong to the group which:

AGREES that:

"Government intervention is good for society"

DISAGREES that:

Identity Priming (Government Dimension Salient)

92
Identity Priming 2 (Government Dimension Salient)
Third-party Allocation Task (2 Salient Dimensions)

Instructions: Allocation Task

Now, you are provided with 100 Credits.
We would like to ask you to allocate the 100 Credits between two randomly selected participants in the study (excluding yourself).

Your decisions will be used to determine the earnings of this randomly selected pair of participants. Likewise, the decisions of some randomly selected participant in the study will be used to determine your earnings.

You will have to make two sets of decisions. One set of your decisions will be randomly selected to be implemented. As before, your decisions will remain anonymous.

Please click next to begin.

Allocation Task Instructions (2 Salient Dimensions)

Allocation Task

Consider the two groups that agree/disagree with the statement

"Government intervention is good for society."

You will now have to make allocations based on which of the two groups each of the two randomly selected participants belongs to.

Allocation Task for Religion Dimension (2 Salient Dimensions)
Allocation Task

Please decide how you would like to allocate the 100 credits between the two randomly selected participants if they belong to the groups below.

“Government intervention is good for society”

Participant who belongs to the group that agrees with the statement
How would I split the $100?
Participant who belongs to the group that disagrees with the statement

0 Credits

Note you can use the arrow keys on the keyboard to adjust the slider precisely.

Allocation for 1 Ingroup 1 Outgroup (2 Salient Dimensions)

Allocation Task

Please decide how you would like to allocate the 100 credits between the two randomly selected participants if they belong to the groups below.

“Government intervention is good for society”

Participant who belongs to the group that agrees with the statement
How would I split the $100?
Participant who belongs to the group that disagrees with the statement

30 Credits

Note you can use the arrow keys on the keyboard to adjust the slider precisely.

Allocation for 2 Ingroups (2 Salient Dimensions)
Allocation Task

Please decide how you would like to allocate the 100 credits between the two randomly selected participants if they belong to the groups below.

**Government intervention is good for society**

- Participant who belongs to the group that disagrees with the statement
- Participant who belongs to the group that agrees with the statement

**Note:** You can use the arrow keys on the keyboard to adjust the slider precisely.

Next

Allocation for 2 Outgroups (2 Salient Dimensions)

Allocation Task

Your allocations to the two randomly chosen participants can either be based on your decisions in the different scenarios earlier, or not. Which would you like?

- Yes, allocate Credits based on the decisions earlier.
- No, allocate Credits **completely** between the two participants.
- No, allocate Credits **exactly** between the two participants.

Next

Allocation Decision for Religion Dimension (2 Salient Dimensions)
**Allocation Task**

Consider the two groups that agree/disagree with the statement

"The world was created by a divine entity."

You will now have to make allocations based on which of the two groups each of the two randomly selected participants belongs to.

---

**Allocation Task for Government Dimension (2 Salient Dimensions)**

Please decide how you would like to allocate the 100 credits between the two randomly selected participants if they belong to the groups below:

"The world was created by a divine entity."

![Diagram showing allocation choices for different groups regarding the statement about divine creation.]

Participants could also be in a group that either agrees or disagrees with the statement "Government intervention is good for society."

Note: You can use the arrow keys on the keyboard to adjust the slider precisely.

---

**Allocation for 1 Ingroup 1 Outgroup (2 Salient Dimensions)**

97
Allocation Task

Please decide how you would like to allocate the 100 credits between the two randomly selected participants if they belong to the groups below:

"The world was created by a divine entity"

Participant who belongs to the group that disagrees with the statement

Participant who belongs to the group that agrees with the statement

How would you split the 100 credits?

0 Credits

0 Credits

Note: You can use the arrow keys on the keyboard to adjust the slider precisely.

Allocation for 2 Ingroups (2 Salient Dimensions)

Allocation Task

Please decide how you would like to allocate the 100 credits between the two randomly selected participants if they belong to the groups below:

"The world was created by a divine entity"

Participant who belongs to the group that agrees with the statement

Participant who belongs to the group that disagrees with the statement

How would you split the 100 credits?

0 Credits

0 Credits

Note: You can use the arrow keys on the keyboard to adjust the slider precisely.

Allocation for 2 Outgroups (2 Salient Dimensions)
Allocation Decision for Government Dimension (2 Salient Dimensions)
Third-party Allocation Task (Religion Dimension Salient)

Instructions: Allocation Task

Now, you are provided with 100 Credits.
We would like to ask you to allocate the 100 Credits between two randomly selected participants in the study (excluding yourself).

Your decisions will be used to determine the earnings of this randomly selected pair of participants. Likewise, the decisions of some randomly selected participant in the study will be used to determine your earnings.

As before, your decisions will remain anonymous.

Please click next to begin.

Allocation Task Instructions (Religion Dimension Salient)

Allocation Task

Consider the two groups that agree/disagree with the statement

“The world was created by a divine entity”.

You will now have to make allocations based on which of the two groups each of the two randomly selected participants belongs to.

Allocation Task (Religion Dimension Salient)
Allocation Task

Please decide how you would like to allocate the 100 credits between the two randomly selected participants if they belong to the groups below.

"The world was created by a divine entity"

Participate who belongs to the group that agrees with the statement

How would I split the 100 credits?

Participate who belongs to the group that disagrees with the statement

$ Credits

You

$ Credits

Note: you can use the arrow keys on the keyboard to adjust the slider precisely.

Allocation for 1 Ingroup 1 Outgroup (Religion Dimension Salient)

Allocation Task

Please decide how you would like to allocate the 100 credits between the two randomly selected participants if they belong to the groups below.

"The world was created by a divine entity"

Participate who belongs to the group that agrees with the statement

How would I split the 100 credits?

Participate who belongs to the group that disagrees with the statement

$ Credits

You

$ Credits

Note: you can use the arrow keys on the keyboard to adjust the slider precisely.

Allocation for 2 Ingroups (Religion Dimension Salient)
Allocation Task

Please decide how you would like to allocate the 100 credits between the two randomly selected participants if they belong to the group below.

"The world was created by a divine entity."

Participant who belongs to the group that disagrees with the statement

How would you split the 100 credits?

Participant who belongs to the group that disagrees with the statement

Note you can use the arrow keys on the keyboard to adjust the slider precisely.

Allocation for 2 Outgroups (Religion Dimension Salient)

Allocation Task

Your allocations to the two randomly chosen participants can either be based on your decisions in the different scenarios earlier, or not. Which would you like?

- Yes, allocate Credits based on the decisions earlier.
- No, allocate Credits randomly between the two participants.
- No, allocate Credits equally between the two participants.

Allocation Decision (Religion Dimension Salient)
Third-party Allocation Task (Government Dimension Salient)

Instructions: Allocation Task

Now, you are provided with 100 Credits.
We would like to ask you to allocate the 100 Credits between two randomly selected participants in the study (excluding yourself).

Your decisions will be used to determine the earnings of this randomly selected pair of participants. Likewise, the decisions of some randomly selected participant in the study will be used to determine your earnings.

As before, your decisions will remain anonymous.

Please click next to begin.

Next

Allocation Task Instructions (Religion Dimension Salient)

Allocation Task

Consider the two groups that agree/disagree with the statement

“Government intervention is good for society”

You will now have to make allocations based on which of the two groups each of the two randomly selected participants belongs to.

Next

Allocation Task (Government Dimension Salient)
Allocation Task

Please decide how you would like to allocate the 100 credits between the two randomly selected participants if they belong to the groups below.

"Government intervention is good for society"

Participant who belongs to the group that disagrees with the statement

Note you can use the arrow keys on the keyboard to adjust the slider precisely.

Allocation for 1 Ingroup 1 Outgroup (Government Dimension Salient)

Allocation Task

Please decide how you would like to allocate the 100 credits between the two randomly selected participants if they belong to the groups below.

"Government intervention is good for society"

Participant who belongs to the group that disagrees with the statement

Allocation for 2 Ingroups (Government Dimension Salient)
Allocation for 2 Outgroups (Government Dimension Salient)

Allocation Task

Please decide how you would like to allocate the 100 credits between the two randomly selected participants if they belong to the group below.

“Government intervention is good for society”

Note: you can use the arrow keys on the keyboard to adjust the slider precisely.

Next

Allocation Decision (Government Dimension Salient)
Identity Elicitation for non-Salient Dimension in 1-Dimension treatment

Now, we would like to ask you several questions about yourself and your experience in the study.

How do you stand on the following issue?
"The world was created by a divine entity"
- Agree
- Disagree

Identity Elicitation for Government Dimension (Religion Dimension Salient)

Now, we would like to ask you several questions about yourself and your experience in the study.

How do you stand on the following issue?
"Government intervention is good for society"
- Agree
- Disagree

Identity Elicitation for Religion Dimension (Government Dimension Salient)
Incentivized Belief Elicitation

Belief Elicitation 1

Belief Elicitation 2
Belief Elicitation 3

Please give your best estimate to the following questions:
the closer your answers are to the true value, the higher will be your earnings in this part.

- One of the four answers will be randomly selected to determine your earnings in this part.
- If your answer selected is equal to the true value (to the nearest integer), you will be given an extra 30 Credits.
- If your answer selected is within 1% of the true value, you will be given an extra 20 Credits.
- If your answer selected is within 2% of the true value, you will be given an extra 10 Credits.
- If your answer selected is not within 2% of the true value, you will not be given any extra Credits.

Q. Consider the statement: "The world was created by a single entity.
What do you think is the proportion of all participants who belong to the group that agrees/disagrees with the statement?"

Belief Elicitation 4

Please give your best estimate to the following questions:
the closer your answers are to the true value, the higher will be your earnings in this part.

- One of the four answers will be randomly selected to determine your earnings in this part.
- If your answer selected is equal to the true value (to the nearest integer), you will be given an extra 30 Credits.
- If your answer selected is within 1% of the true value, you will be given an extra 20 Credits.
- If your answer selected is within 2% of the true value, you will be given an extra 10 Credits.
- If your answer selected is not within 2% of the true value, you will not be given any extra Credits.

Q. Consider the statement: "Government intervention is good for society.
What do you think is the proportion of all participants who belong to the group that agrees/disagrees with the statement?"
Identity Strength Elicitation

Group that disagrees with the statement "The world was created by a divine entity".

To what extent do you see yourself as belonging to the group above? Please indicate using the slider below.

Scale Identity Strength for Religion Dimension

Group that agrees with the statement "Government intervention is good for society".

To what extent do you see yourself as belonging to the group above? Please indicate using the slider below.

Scale Identity Strength for Government Dimension
Post-Experiment Survey Questions

Questions related to decision task

Did you understand the instructions?  [Did understand., Did not understand.]

Please write here if there is anything unclear about the study

In the allocation task, how did you make the decisions? (Select the option which is most important to you.)
[I made the decisions based on my preferences., I made the decisions based on what I wanted to do., I made the decisions based on what I thought the experimenter expected me to do., I made the decisions randomly., I made the decisions based on gut instinct., I made the decisions based on what I thought was the socially desired thing to do., I made the decisions based on what I thought was fair.*, Other: Please describe in the textbox below.]

Demographic questions

In which year were you born?

What is your gender?
[Male, Female, I would prefer not to say, Other (Please describe below if you wish)]

What is your highest education level completed?
[Less than High School Diploma, High School Diploma, Associate’s Degree (AA/AS), Bachelor’s Degree (BA/BS), Master’s Degree (MA/MS/M.Eng), Doctoral Degree (PhD), Professional Degree (e.g., JD, MD)]

Which state in the US are you currently living in?

What is your annual personal income?
[Under $10,000, $10,000 to $19,999, $20,000 to $29,999, $30,000 to $39,999, $40,000 to $49,999, $50,000 to $59,999, $60,000 to $69,999, $70,000 to $79,999, $80,000 to $89,999, $90,000 to $99,999, $100,000 to $124,999, $125,000 to $149,999, $150,000 and above]

Consider the total income brought into your household by all members including yourself. What percentage of this total household income is made up of your income?
[None - 0%, 0-10%, 10-20%, 20-30%, 30-40%, 40-50%, 50-60%, 60-70%, 70-80%, 80-90%, 90-100%, All - 100%]

What is your ethnicity?
[White, Black or African American, Native American or American Indian, Hispanic or Latino, Asian or Pacific Islander, Others or Multiracial]

In general, what is your political affiliation?
[Democrat, Republican, Independent, Other, None]

What is your religious affiliation?
[Baha’i, Buddhism, Candomble, Christianity (e.g. Baptist, Church of England, Roman Catholic, Methodist, Jehovah Witness, etc.), Hinduism, Islam, Jainism, Judaism, Non Religious (e.g. Agnostic, Atheist, No Religion), Paganism, Rastafari, Santeria, Shinto, Sikhism, Spiritualism, Taoism, Unitarianism, Zoroastrianism, Other, Do Not Wish to Answer]

*This option only appeared in the third session.
Appendix C (for Chapter 3)

C.1 Information Preference Scale

In each scenario below, you will have an opportunity to receive information. This information may or may not be useful and it may or may not be painful to learn. Please read each scenario carefully, then indicate if you want to know that information. [Choices: Definitely don’t want to know; Probably don’t want to know; Probably want to know; Definitely want to know. “R” is scored in reverse.]

1) As part of a semiannual medical checkup, your doctor asks you a series of questions. The answers to these questions can be used to estimate your life expectancy (the age you are predicted to live to). Do you want to know how long you can expect to live?

2) You provide some genetic material to a testing service to learn more about your ancestors. You are then told that the same test can, at no additional cost, tell you whether you have an elevated risk of developing Alzheimer’s. Do you want to know whether you have a high risk of developing Alzheimer’s?

3) At your annual checkup, you are given the option to see the results of a diagnostic test, which can identify, among other things, the extent to which your body has suffered long-term effects from stress. Do you want to know how much lasting damage your body has suffered from stress?

4) Ten years ago, you had the opportunity to invest in two retirement funds: Fund A and Fund B. For the past 10 years, you have invested all your retirement savings in Fund A. Do you want to know the balance you would have if you had invested in Fund B instead?

5) You decide to go to the theater for your birthday and give your close friend (or partner) your credit card so they can purchase tickets for the two of you, which they do. You aren’t sure but suspect that the tickets may have been expensive. Do you want to know how much the tickets cost?

6) You bought an electronic appliance at a store at what seemed like a reasonable, though not particularly low, price. A month has passed, and the item is no longer returnable. You see the same appliance displayed in another store with a sign announcing “SALE.” Do you want to know the price you could have bought it for?

7) You gave a close friend one of your favorite books for her birthday. Visiting her apartment a couple of months later, you notice the book on her shelf. She never said anything about it; do you want to know if she liked the book?

8) Someone has described you as quirky, which could be interpreted in a positive or negative sense. Do you want to know which interpretation he intended?
9) You gave a toast at your best friend’s wedding. Your friend says you did a good job, but you aren’t sure if he or she meant it. Later, you overhear people discussing the toasts. Do you want to know what people really thought of your toast?

10) As part of a fundraising event, you agree to post a picture of yourself and have people guess your age (the closer they get, the more they win). At the end of the event, you have the option to see people’s guesses. Do you want to learn how old people guessed that you are?

11) You have just participated in a psychological study in which all of the participants rate others’ attractiveness. The experimenter gives you an option to see the results for how people rated you. Do you want to know how attractive other people think you are?

12) Some people seek out information even when it might be painful. Others avoid getting information that they suspect might be painful, even if it could be useful. How would you describe yourself?

13) If people know bad things about my life that I don’t know, I would prefer not to be told. [R]
C.2 Five Facet Mindfulness Questionnaire

Please indicate how true the below statements are of you using the scale provided. [Choices: Never or very rarely true; Rarely true; Sometimes true; Often true; Very often or always true. “R” is scored in reverse. Observing items: 1, 6, 11. Describing items: 2, 7, 12. Acting with awareness items: 3, 8, 13. Non-judging items: 4, 9, 14. Non-reacting items: 5, 10, 15.]

1) When I take a shower or a bath, I stay alert to the sensations of water on my body.

2) I’m good at finding words to describe my feelings.

3) I don’t pay attention to what I’m doing because I’m daydreaming, worrying, or otherwise distracted. [R]

4) I believe some of my thoughts are abnormal or bad and I shouldn’t think that way. [R]

5) When I have distressing thoughts or images, I “step back” and am aware of the thought or image without getting taken over by it.

6) I notice how foods and drinks affect my thoughts, bodily sensations, and emotions.

7) I have trouble thinking of the right words to express how I feel about things. [R]

8) I do jobs or tasks automatically without being aware of what I’m doing. [R]

9) I think some of my emotions are bad or inappropriate and I shouldn’t feel them. [R]

10) When I have distressing thoughts or images I am able just to notice them without reacting.

11) I pay attention to sensations, such as the wind in my hair or sun on my face.

12) Even when I’m feeling terribly upset I can find a way to put it into words.

13) I find myself doing things without paying attention. [R]

14) I tell myself I shouldn’t be feeling the way I’m feeling. [R]

15) When I have distressing thoughts or images I just notice them and let them go.
C.3 Perceived Stress Scale

The questions below ask about your feelings and thoughts during the last week. For each question, you will be asked to indicate how often you felt or thought a certain way. Although some of the questions are similar, there are differences between them and you should treat each one as a separate question. The best approach is to answer each question fairly quickly. That is, don’t try to count up the number of times you felt a particular way, but rather indicate the alternative that seems like a reasonable estimate. [Choices: Never; Almost never; Sometimes; Fairly often; Very often. “R” is scored in reverse.]

1) In the last week, how often have you been upset because of something that happened unexpectedly?

2) In the last week, how often have you felt that you were unable to control the important things in your life?

3) In the last week, how often have you felt nervous and stressed?

4) In the last week, how often have you felt confident about your ability to handle your personal problems? [R]

5) In the last week, how often have you felt that things were going your way? [R]

6) In the last week, how often have you found that you could not cope with all the things that you had to do?

7) In the last week, how often have you been able to control irritations in your life? [R]

8) In the last week, how often have you felt that you were on top of things? [R]

9) In the last week, how often have you been angered because of things that happened that were outside of your control?

10) In the last week, how often have you felt difficulties were piling up so high that you could not overcome them?
### C.4 Appendix Figures and Tables

Table C.4.1: Baseline Measures

<table>
<thead>
<tr>
<th>Variables</th>
<th>All Mean</th>
<th>All SD</th>
<th>Treatment Mean</th>
<th>Treatment SD</th>
<th>Control Mean</th>
<th>Control SD</th>
<th>Diff Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age(^a)</td>
<td>43.81</td>
<td>12.61</td>
<td>43.81</td>
<td>11.73</td>
<td>43.81</td>
<td>13.47</td>
<td>0.00</td>
</tr>
<tr>
<td>Female</td>
<td>0.61</td>
<td>0.49</td>
<td>0.60</td>
<td>0.49</td>
<td>0.61</td>
<td>0.49</td>
<td>0.00</td>
</tr>
<tr>
<td>White</td>
<td>0.93</td>
<td>0.25</td>
<td>0.92</td>
<td>0.27</td>
<td>0.94</td>
<td>0.24</td>
<td>-0.01</td>
</tr>
<tr>
<td>Degree</td>
<td>0.57</td>
<td>0.50</td>
<td>0.55</td>
<td>0.50</td>
<td>0.60</td>
<td>0.49</td>
<td>-0.05</td>
</tr>
<tr>
<td>Household income (1-10)</td>
<td>4.63</td>
<td>2.30</td>
<td>4.77</td>
<td>2.31</td>
<td>4.49</td>
<td>2.29</td>
<td>0.28</td>
</tr>
<tr>
<td>Conservatism (0-100)</td>
<td>44.17</td>
<td>22.07</td>
<td>45.57</td>
<td>22.52</td>
<td>42.75</td>
<td>21.60</td>
<td>2.82</td>
</tr>
<tr>
<td>Perceived stress (0-40)</td>
<td>17.84</td>
<td>3.99</td>
<td>17.78</td>
<td>4.05</td>
<td>17.90</td>
<td>3.94</td>
<td>-0.12</td>
</tr>
</tbody>
</table>

Observations: 261 / 131 / 130

**Notes:** None of the differences in means were significant at the 10% level. “Degree” is whether they have a Bachelor’s degree. “Household income” bracket \( i \) is \((i-1)\times£10,000 \) to \( i\times£10,000 \) (pre-tax). “Conservatism” is liberal-conservative scale. “Two participants in the treatment group did not give their age, so the number of observations on age in the full sample / treatment was 259 / 129.”
Figure C.4.1: Number of Participants at Each Session

![Graph showing the number of participants at each session with two lines, one for control and one for treatment. The graph has a y-axis labeled "Number of Participants" ranging from 105 to 135 and an x-axis with days from Pre-course to Post-course. The graph illustrates a decrease in participants over time for both groups.]
Table C.4.2: Comparison of Baseline Characteristics of Non-Attritors in Treatment and Control

<table>
<thead>
<tr>
<th></th>
<th>Difference in Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.118 (1.695)</td>
</tr>
<tr>
<td>Female</td>
<td>-0.041 (0.065)</td>
</tr>
<tr>
<td>White</td>
<td>-0.013 (0.036)</td>
</tr>
<tr>
<td>Degree</td>
<td>-0.093 (0.065)</td>
</tr>
<tr>
<td>Household income (1-10)</td>
<td>0.337 (0.310)</td>
</tr>
<tr>
<td>Conservatism (0-100)</td>
<td>2.956 (2.953)</td>
</tr>
<tr>
<td>Perceived stress (0-40)</td>
<td>-0.116 (0.523)</td>
</tr>
</tbody>
</table>

*Notes:* Standard errors in parentheses. No differences were significant at the 10% level.
Figure C.4.2: Average Session Evaluation

Notes: This figure shows the distributions over participants of their average session evaluation during the interventions (bins are of width 1). The distributions are similar for the treatment and control; average feedback per session was 3.92 in the control and 3.75 in the treatment.
Figure C.4.3: Compliance Measure (Average Browser Tab Switches per Session)

Notes: This figure shows the distributions over participants of the average number of browser tab switches per session during the interventions (bins are of width 1). The distributions are similar for the treatment and control, with significant numbers of participants focusing during the recordings (over 40% of the treatment and control groups have an average number of switches between 0 and 1). The difference in the mean of the control (1.61) and the treatment (1.72) is not statistically significant ($p = 0.630$).

Figure C.4.4: Compliance Measure (Average Difference in Submission Time per Session)

Notes: This figure shows the distributions over participants of their average difference in submission time per session from the true end of the recording (bins have a width of 30 seconds). The distributions are similar for the treatment and control, with a substantial portion of participants (over 60%) in both groups submitting more or less when they are told to (within 30 seconds of the end of the recording). The difference in the mean of the treatment (58 seconds) and control (36 seconds) is not statistically significant ($p = 0.143$).
Figure C.4.5: Intervention Effects on the Five Facets of Mindfulness Scale

**Notes:** This figure compares the post-course means of the FFMQ scale and sub-scales in control and treatment group. Gray bars show 95% confidence intervals.


Everett, J. A. (2013). The 12 item social and economic conservatism scale (sees). *PloS one* 8(12), e82131.


