Polarisation, personality, and discriminatory behaviour: Experiments in Zimbabwe and the United Kingdom.

Lerato Dixon

A thesis submitted in partial fulfilment of the requirements of Nottingham Trent Universityfor the degree of Doctor of Philosophy

Supervisors: Professor Thorsten Chmura, Professor Robert Ackrill and Dr Jingwen Fan

August 2022

Copyright statement

Abstract

The thesis presents four main research projects using behavioural and experimental economics methods. Chapter 2 presents an experiment that investigates discriminatory behaviour in one-shot experimental games conducted in Zimbabwe. We also use social norm-based interventions to shift behaviours. Chapter 3 reports an experiment replicating methodology with context-specific variations used in the previous chapter but conducted in the United Kingdom. In both experiments. The results present evidence of out-group discrimination. Salience or association with the identity also influences discriminatory behaviour. Beliefs about others' cooperativeness drives their contributions. The social norm-based nudges have varied effects on behaviours, at times having the opposite effect than intended. However, the treatment effectively alters modal responses in the social norms elicitation task. Chapter 4 investigates the impact of social identities on their willingness to lie for others. Using a variation of the Mind Game, individuals have the opportunity to lie on behalf of someone else. Our findings are consistent with the literature that participants do not maximise their payoffs over ten rounds in both the United Kingdom and Zimbabwe experiments. We show that salience to a group identity influences the likelihood, emphasising the moral cost of dishonesty. We also consider normative acceptance of bribery. This paper also contributes to the experimental literature on corruption and dishonest behaviour using a variation of the Mind Game. Lastly, chapter 5 examines individual-level personality traits using the Big Five Inventory. We compare personality traits between Leave and Remain voters in the United Kingdom. We also investigate the influence of personality traits on discriminatory behaviour. The paper also briefly considers the potential for misrepresenting political identities and the potential relationship between personality traits. Our results suggest that Leavers and Remainers differ in Openness, with Remainers reporting higher levels of Openness. In addition, we find evidence that, to some extent, personality traits dictate cooperative behaviours. The findings contribute to the research on personality traits, political decisions and behaviours.

Acknowledgements

Firstly, to my supervisory team, Thorsten, Rob and Jingwen. Thorsten, for continually challenging and encouraging me. Thank you for the excellent support, guidance and advice. Rob, for always bringing up new ideas and insights to think and mull over. Jingwen, for all of the advice. I'm grateful for all the guidance, wisdom and support. I cannot thank you enough.

I would like to acknowledge the Economics department at NBS, where I am currently a faculty member, for the financial support in funding my research and supporting my journey.

Secondly, to the colleagues, mentors and friends that have guided and taught me over this time. As well as members of the Centre for Behavioural Sciences (CBS) for suggestions and useful input. Amairisa, I'm grateful for the discussions, support and laughter. My gratitude also extends to the assistance and help from Nyadire University, Harare Polytechnic, Bulawayo Polytechnic and Marondera University.

I'm grateful for the continued support from friends. Who have prayed for me, encouraged me and laughed with me.

A special thank you goes to my family. Although I could not share this journey with my father, Dumisani, I am grateful for the interest in the world you instilled in me. To my mother, Pauline, I am forever grateful that you did not let tragedy turn into hell because of your strength and perseverance.

Finally, a special thank you to my husband, Matt, for his continual support and encouragement. You gave me the time and space to work on my PhD. Thank you for your unwavering belief in me and unconditional love. Not forgetting the sweetest and craziest dog, Akeema, who brings lots of joy and laughter.

Contents

List of Figures

List of Tables

1

Introduction

1.1 General introduction

This thesis presents four research chapters that are embedded in the field of behavioural and experimental economics. Each chapter is a stand-alone research chapter that can be read separately. The four chapters have shared themes and similarities in objective and methodology to address the research questions. The overarching aim of this research is to analyse intra-group favouritism and inter-group discriminatory behaviour in two empirical contexts; and the extent to which social norm-based can shift behaviours between groups.

All four chapters examine the effect of social identity in various settings to understand the impact on economic interactions between individuals. Contrary to standard economic theory, laboratory and field experiments have evidenced that individuals do not simply act out of self-interest but rather consider cooperativeness, trust and their broader identity. In the 1950s, economists began considering the role of discrimination in markets. The literature differentiates between two models of discrimination: taste-based and statistical discrimination. Taste-based discrimination describes when agents gain utility from favouring

their own group over another Becker (1971). Statistical discrimination arises when agents draw upon incomplete information about group characteristics to dictate how to behave (Arrow 1972).

In-group favouritism and out-group discrimination depart from traditional economic theory. Various models and theories have been identified for deviations from the standard model. One such model is Social Identity Theory (SIT). The SIT posits that people either maximise in-group outcomes or maximise differences between in-group and out-group outcomes, even in minimal or artificial groups - exogenously induced by the experimenter - (Tajfel et al. 1971). Agents identify themselves based on mere self-categorisation into different groups separate from their personal or individual traits. Individuals will favour their in-group at the expense of the out-group (Tajfel et al. 1971, Taijfel 1970, Turner et al. 1979). Likewise, Akerlof & Kranton (2000) stress the importance of identity in influencing economic behaviour. The theory posits that economic agents gain utility from behaving appropriately with respect to their social group. Identifying with a specific group will impact behaviour with in-group members and out-group members (Chmura et al. 2005). This identity is internal to agents; there may be the threat of sanction if one deviates from the group's norms, even when alone. The identity they prescribe themselves and the associated social norms guide individuals towards certain behaviour, even at the expense of individual interest (Elster 1989).

An important consideration when investigating group identity is the opportunity for individuals to misreport their identity. Inconsistent labelling by participants has the potential to undermine identity-based research. Therefore, in chapter 4, we also analyse the identities reported to the experimenter versus the recruitment platform.

An extension of SIT is the concept of polarisation, which can also explain differences in behaviours between groups of individuals (Cason & Mui 1997, Mackie 1986, Luhan et al. 2009). In general, polarisation is characterised by homogeneity within groups, heterogeneity between groups and alternative groups having an insignificant effect (Esteban & Ray 1994, Sunstein 2019*a*, McCoy et al. 2018, Druckman et al. 2021, Iyengar & Westwood 2015). Moreover, an identity group needs to be salient to the individual (Böhm et al. 2020). The research combines experimental findings with self-reported polarisation measures to identify how polarisation has occurred and affects discriminatory behaviour (Sunstein 1999, 2019*a*). The experimental method has allowed for a quantifiable measure of in-group favouritism and out-group discrimination contingent on primed identities. In two of our chapters, we briefly consider group polarisation and how it relates to the political contexts and discriminatory behaviour.

Another leading explanation for socially motivated behaviours is social norms. Social norms, to an extent, govern behaviour in different environments and deviating from this behaviour can lead to punishment of various forms. Social norms prescribe or proscribe behaviours or actions that are collectively perceived (Ostrom 1998, Elster 1989, Krupka & Weber 2013, Cialdini et al. 1991, Ostrom 2014). Therefore, self-image concerns, fear of sanction, or the intrinsic costs of violating social norms impacts behaviours. The thesis also contributes to the literature exploring the influence of social norms on behaviour and the effectiveness of social norm-based interventions to shift behaviours (Thaler & Sunstein 2009). In chapters two and three, novel norm-based interventions are implemented to evaluate how behaviour can be shifted (Bicchieri 2015, Dimant et al. 2020, Bolton et al. 2019). Therefore, we exogenously manipulate the prevailing norms using two social norms: *Heal*, which we hypothesise will foster greater cohesion between groups and *Divide*, which we hypothesise will further separate groups. We expand on similar work by Bursztyn et al. (2020); however, in our study, we include two novel manipulations rather than one.

The third chapter examines group relations and social image concerns in a lying game. Individuals are regularly faced with environments with opportunities for dishonest behaviour and opportunities where dishonesty can benefit others. Again, contrary to standard economic theory, agents do not behave entirely dishonestly. It has been posed that social ties may increase the likelihood of dishonesty as it can be seen to benefit the group (Wiltermuth 2011, Feldhaus & Mans 2014). Conversely, social identity and norms may keep people accountable as they aspire to be seen to be doing the right thing to preserve their social image (Dai et al. 2018, Mazar et al. 2008, Cohn et al. 2019). In two experiments conducted in the United Kingdom and Zimbabwe, we test people's willingness to lie for themselves and others and the impact of social norm concerns on dishonest behaviour.

The fourth research chapter builds on these lessons by considering how underlying individual-level personality traits dictate behaviours that motivate social preferences and voting outcomes. In this chapter, we look in more detail at the personality traits at play in cooperative behaviour.

At this point, it is worth noting the impact of COVID-19 on the primary data collection for the thesis. Data collected in Zimbabwe was stopped prematurely due to the global lockdown. The subsequent effects are that the sample size collected in Zimbabwe is below what was initially set out; this also impacts numbers within the various treatments. Despite the halting data collection, we are able to glean some insights into behaviours across various contexts. In addition, data collection was then switched from laboratory in-person sessions

to online experiments via Prolific Recruitment service; hence, all data collection for the research focused on the United Kingdom occurred using online experiments.

1.2 Outline of thesis

Chapter two, titled "Polarisation and discrimination in Zimbabwe". Zimbabwe provides a pertinent case of polarisation and discrimination. Based on deep-rooted ethnic cleavages, particularly following independence and onwards, there was rhetoric in the political sphere that identified a clear binary divide around supporting Zimbabwe versus being against Zimbabwe and a traitor (LeBas & Munemo 2019). We find evidence of out-group discrimination and in-group favouritism. Further still, group salience also influences inter-group behaviours. We also observe that social norms significantly impact behaviour across all game forms. Following the norm-based interventions in the second stage of the experiment, we find that the norms significantly affect cooperation, altruism, and trust. In general, the *Heal* treatment fosters greater prosocial behaviour; however, it backfires with Shona participants and leads to lower in-group cooperation. Contrary to our hypothesis, Ndebele participants in the *Divide* treatment increases cooperation for all receivers. However, for Shona participants in the *Divide* treatment group, we find further division with the Zimbabwean receiver. In terms of shifting the evaluation of social norms, we find that the *Divide* intervention significantly impacts social norm evaluations. This chapter extends the experimental literature on inter-group behaviours in developing countries.

Chapter three is titled "Discrimination in the UK: Leavers and Remainers". Following the 2016 referendum on the United Kingdoms European Union membership, a dichotomous divide was established as being a Leaver or a Remainer, forming a salient opinion-based identity (Hobolt et al. 2021, Hobolt 2016). The results from the experiments reveal that Leavers contribute more to fellow Leavers. Regarding the polarisation measures, participants who felt like a 'winner' following the Referendum are less prosocial, leading to out-group discrimination. Similarly, participants who report that they would vote differently given another Referendum are less prosocial. Surprisingly, we find that the evaluations of social norms do not significantly impact game behaviour, perhaps due to the newness of this opinion-based political identity group (Druckman et al. 2021). We find that the norm-based interventions impact behaviours in experiments, particularly in group settings (more than two players), such as the Public Goods Game. In particular, both treatments diminish in-group cooperativeness for Remainers. Conversely, both treatments bridge the gap between

Leavers and Remainers. Findings from chapters two and three point to the importance of salient naturally-occurring identities, which has been debated in the literature with mixed results from ethnic, political and artificial groups. The thesis also contributes to the literature exploring the influence of social norms on behaviour and the effectiveness of social norm-based interventions to shift behaviours in two diverse yet polarised groups.

The following chapter, entitled "Would I lie for you", investigates the extent that people are willing to lie on behalf of someone else. Earlier experimental studies investigating dishonesty find that individuals partially lie, failing to maximise their material payoff. One of the leading explanations in the experimental literature for why individuals lie partially is that agents have an aversion to lying. Individuals will prefer to avoid being dishonest, even if being dishonest yields a positive payoff. Therefore, individuals balance the payoff from dishonest behaviour versus the disutility from lying (Erat & Gneezy 2012, Gneezy et al. 2018, Khalmetski & Sliwka 2019, Mazar et al. 2008). Additionally, social image concerns may be a principal determining factor in mitigating complete dishonest behaviour and thus can explain partial dishonest behaviour (Bernhard et al. 2006, Fischbacher & Föllmi-Heusi 2013, Gächter & Schulz 2016, Grolleau et al. 2016).

This chapter explores honesty using a novel variation of the Mind Game by Kajackaite & Gneezy (2017) to understand further the propensity to lie for others when it also benefits the liar. In this chapter, participants are randomly allocated into treatment groups, where the identity of the co-benefiter of dishonesty is revealed. In both experiments, we find that individuals lie for themselves. We find that in Zimbabwe, it is not the reported identity but the degree to which the individual aligns with that identity that affects their propensity for dishonesty for others and themselves. Therefore, we observe a moral cost of dishonest behaviour when engaging with someone within a group when the individual identifies high salience with the group identity. Lastly, we find evidence that norms around actions corresponding to bribery impact dishonesty. This chapter contributes to the experimental literature on corruption and dishonest behaviour using a novel variation of the Mind Game.

chapter 5 is entitled "Voting outcomes and the big five personalities". In this chapter, we investigate whether voting decision and cooperativeness is correlated with certain personality traits. These personalities are grouped into five taxonomies: Extroversion, Agreeableness, Emotional Stability, Conscientiousness, and Openness (John et al. 2008). Previous studies have identified that the Big Five Inventory (BFI) predicts behaviours in various environments, from sending in the dictator game to group action in political participation (Ben-Ner et al. 2004, Robalo et al. 2017). We measure personality characteristics

using a variation of the BFI as outlined by John et al. (2008). This chapter uses 18 short sentences on prototypical trait indicators of the Big Five taxonomies. We observe that Remainers are more open than Leavers. Our results also show that personality traits play a role in taste-based discrimination. Specifically, we find evidence that individual characteristics such as agreeableness and extraversion significantly affect contributions across the game forms. Lastly, extraversion and conscientiousness play a role in misreporting. In particular, an interesting finding is that conscientious Leave voters (Remain voters reported to the recruitment platform) are more likely to have misreported their voting decision between the experimenter and recruitment platform, perhaps due to cautiousness of the negative connotations. Our findings provide methodological contributions to the experimental literature and further develop an understanding of personalities, voting and group behaviours. Further implications are discussed in the chapter.

Chapter six concludes the thesis by reviewing the results from the four chapters. We identify any limitations and provide suggestions for future meaningful research.

2

Polarisation and discrimination in Zimbabwe

Abstract

We experimentally test polarisation and discriminatory behaviour in Zimbabwe. From the perspective of Social Identity Theory, we investigate how ethnolinguistic identities in Zimbabwe. To identify discriminatory behaviour, we vary the ethnicity of the receivers. We also include social norm-based treatments. Using a proscriptive social norm, we test if discriminatory behaviour can be further exacerbated. In the alternative treatment, we use a prescriptive social norm nudge to test if discriminatory behaviour can be minimised. We also investigate the relationship between discriminatory behaviour, beliefs and social norms. We find evidence of discriminatory behaviour, particularly concerning salient identities. Beliefs largely determine behaviours. We observe treatment effects following the norm-nudge treatments. Both the proscriptive (*Divide*) and prescriptive (*Heal*) treatments yield more robust results, but the strength of the effect varies across different choice contexts.

Keywords: Discrimination; Polarisation; Identity; Social norm-nudge; Zimbabwe; Social norms

2.1 Introduction

Group polarisation and discrimination between groups can have negative economic and social consequences (Ostrom 1998, Schaub 2017). Easterly & Levine (1997) find that ethnic division negatively impacts economic development. As ethnic groups move away from one another, with solidarity with the in-group, which is opposed to the out-group. In the 1950s, economists began considering the role of discrimination in markets. The literature started to differentiate between two models of discrimination: taste-based and statistical discrimination. *Taste-based discrimination* arises when agents gain utility from favouring their group over another, even willing to pay a financial cost to avoid the out-group (Becker 1971). *Statistical discrimination* arises when incomplete information is drawn upon to characterise a group (Arrow 1972, Phelps 1972). These two forms of discrimination support findings observed in economic studies (Abbink & Harris 2012, Guryan & Charles 2013, Akerlof & Kranton 2000, 2005). We conduct our experiments on naturally-occurring groups in Zimbabwe. Specifically, the deep-rooted divisions between the two ethnolinguistic groups in Zimbabwe, the Shona and Ndebele, which provide a pertinent example of in-group favouritism and out-group discrimination.

In Zimbabwe, racial identities are synonymous with the language spoken (Ndhlovu 2007). The main indigenous languages spoken are Shona and Ndebele. The majority race is Shona, who constitute approximately 70% of the Zimbabwean population (Comparative Constitution Project 2018), and the significant minority race is the Ndebele, that constitutes approximately 20% of the Zimbabwean population. Historically, there have been mixed narratives and contrasting accounts of the Zimbabwean identity (Ndlovu-Gatsheni 2012). After British colonialism, ethnic differences were emphasised. African nationalism created and changed identities, fragmented and unified different groups (Ndlovu-Gatsheni 2012). The Rhodesian government stressed race through various measures such as Identity Cards that detailed village and place of origin; these ID cards remain the same post-independence. The use of identity cards has been significant in enforcing historical notions of distinct identities between the Ndebele and Shona (Muzondidya & Ndlovu-Gatsheni 2007). Moreover, where someone is from is often asked in Zimbabwe and is often used to identify the language, customs and even political point of view of someone (Muzondidya & Ndlovu-Gatsheni 2007, Macy et al. 2019).

Polarisation and discrimination have been categorised in the psychological literature as having their origins in social identity theory (Tajfel et al. 1971, Taijfel 1970, Turner et al. 1979).

Social identity theory posits that cooperation and trust are more likely amongst members with the same identity because of in-group favouritism (Tajfel et al. 1971). There has been substantial experimental research on the effects of ethnicity, race and nationality on prosocial behaviour, a benefit to another individual or group of people at a cost to the self (Robinson 2016, Chuah et al. 2007, Fong & Luttmer 2011, Burns 2004). For instance, research on ethnic favouritism has shown that this fosters political favouritism and has been found to benefit the areas in which political leaders are from or share ethnic identities with (De Luca et al. 2018, Posner 2005). The experimental literature has tested ethnic division using a variety of economic games to study in-group favouritism and out-group discrimination, typically using a small variety of experimental games with selected modifications. Typically, the experimenters reveal the identity of the other participant, using surnames, images, or explicitly informing participants of the identity of the other player, then measure differences in behaviour between those of the in-group who share the specific identity characteristics and the out-group who do not share these specific identity characteristics (Ahmed 2010, Der Merwe & Burns 2008, Grosskopf & Pearce 2017). The experimental method allows for the degree of economic and social efficiency to be measurable and quantifiable.

In addition, scholars have identified that the prevailing social norms surrounding discriminatory behaviour are contingent on a shared identity group and should be closely related to realised discriminatory behaviour (Barr et al. 2018, Ostrom 2014, 1998, Falk & Zehnder 2013). Norms are not constantly prevalent in individual decision-making; therefore, experimentalists need to ensure that these norms are drawn upon in the experimental procedure (Krupka & Weber 2013, Smerdon et al. 2020, D' Adda et al. 2016, Bicchieri & Dimant 2019). In our study, individuals participated in a modified version of the social norms elicitation task by Krupka & Weber (2013). In our social norms elicitation task, we asked participants to evaluate the social appropriateness of various behaviours in different real-world scenarios. Hence, this paper also considers the role of social norms in influencing discrimination or favouritism in Zimbabwe. Therefore, we assess the relationship between social norms and incidences of favouritism and discrimination.

This paper also attempts to mitigate the impacts of division using behaviourally informed interventions. Behavioural public policy and behavioural public administration (Ewert & Loer 2021, Gofen et al. 2021) have received increased attention (Bicchieri 2015, Ciriolo et al. 2019). Bicchieri (2015) find that norm-based interventions can alter individuals' behaviour when considering gender roles. In general, people form beliefs about others to explain divisions between groups. Social norms, to an extent, govern behaviour in different environments

and deviating from this behaviour can lead to punishment of various forms. Thus, social norms and the framing around social norms can be effective policy tools (Cialdini et al. 2006, Miller & Prentice 2016).

Our experiments are comprised of two main parts. Having asked participants about their ethnicity, the first part of the experiments involved participants undertaking a variety of within-subject experimental games using well-documented experiments in the literature - the prisoner's dilemma, dictator game, trust game and public goods game. In each game, we reveal the receiver's identity as Ndebele, Shona and Zimbabwean citizen. In other words, participants have to distribute their endowment to an *in-group*, *out-group* and *Mixed-group* member[1]. Following the games, individuals participate in a modified version of the social norms elicitation task by Krupka & Weber (2013). In our social norms elicitation task, we asked participants to evaluate the social appropriateness of various behaviours in different real-world scenarios; their own perceptions of the social appropriateness of actions, the majority of in-group perceptions of the socially appropriate action, and the majority of out-group perceptions of the socially appropriate action. Lastly, participants complete a post-experiment questionnaire to identify demographic characteristics and self-reported identity salience as a quantitative measure of polarisation.

In the second part of the experiment, participants are randomly assigned to one of the two norm-based treatments and a control group who receive no information to evaluate how behaviour can be shifted. We exogenously manipulate the prevailing norms before undertaking the games, using two social norms: *Heal*, a prescriptive social norm (Farrow et al. 2017) which details what people approve of doing, which we hypothesise will foster greater cohesion between the two ethnicities and *Divide*, a proscriptive social norm which details what people disapprove of doing, which we hypothesise will further separate the two ethnicities. We expand on similar work by Bursztyn et al. (2020); however, in our study, we include two novel manipulations rather than one. Therefore, in light of the polarisation and division in Zimbabwe, we are also interested in investigating if differences between the Shona and Ndebele can be bridged and how the division can be further fractured.

The analysis of identity, polarisation and discriminatory behaviour is consistent with the literature. In-group favouritism and out-group discrimination is observed. In addition, elicited beliefs strongly affect behaviour (Bigoni et al. 2019, Fischbacher & Gächter 2010). We find that the self-reported measures of polarisation, such as individuals with more substantial

[1]Participants are told that the Zimbabwean receiver has a 50% probability that the individual is Shona or Ndebele.

alignment to their ethnic group, behave less cooperatively with out-group members. These finding mirrors findings by Mill & Morgan (2021). In addition, we find that the norm treatments have mixed effects on behaviour, with varying directional impacts.

This paper contributes to the behavioural literature by identifying the relationship between discriminatory behaviour, social norms and the degree to which norms can be utilised as practical policy tools. We also contribute to the growing incorporation of experimental and behavioural economic methods and models within policy. Thirdly, to the best of our knowledge, we contribute to the experimental discriminatory behaviour literature by using Ndebele participants who have not been studied in the behavioural literature. We expand on the pre-existing literature investigating the impact of naturally occurring identity groups on social preferences and cooperation (Abbink & Harris 2012, Ahmed 2010, Carpenter & Cardenas 2011, Chuah et al. 2013, Etang et al. 2011, Kim et al. 2017, Slonim & Guillen 2010).

The remainder of this paper is organised as follows. In section 2.2, we briefly discuss the relevant literature and further background. In section 2.3, we identify the theoretical framework and hypotheses. In the following section, we present our experimental design and procedure. In section 2.5, we detail the results of the experiments. Lastly, in Section 2.6, we summarise the study, identify some limitations, and provide recommendations for further research.

2.2 Background and related literature

2.2.1 The Zimbabwean context

Since gaining independence in 1980, although being a nationalist group, the Zimbabwean state and ruling party has often appeared to favour Shona communities at times at the expense of other minority communities (Ndlovu-Gatsheni 2011). Partly because, since independence, ruling politicians have been Shona-speakers (LeBas & Munemo 2019). As the political party that led to independence, ZANU, may have successfully built an independent state but failed to create a unified state (Ndlovu-Gatsheni 2012). Since the 1980's political conflict has been tied to ethnic groups (CCJP 1997), political parties became recognised as Ndebele-speaking and Shona-speaking; different ethnic political leaders were put against each other (Muzondidya & Ndlovu-Gatsheni 2007, LeBas & Munemo 2019). Following independence, President Robert Mugabe ordered the Fifth Brigade (a Shona armed force) to

kill, rape and pillage areas that supported the opposing party, ZAPU. The ZAPU areas were mainly within Matabeleland (mostly Ndebele areas within Zimbabwe). The explicit ethnic targeting perpetuated differences between Shona and Ndebele and encouraged more radical politics from both political parties (Muzondidya & Ndlovu-Gatsheni 2007).

ZANU-PF have often used and baited the national identity for their own purposes (Onslow 2011). Research by Posner (2005) observed that ethnic identities signal an individuals voting preferences and can have an effect in both political and non-political contexts. As noted in Zambia, most political contests are based around ethnic or tribal divisions (Posner 2005). Moreover, democratic practices such as referenda can be used to further division and polarisation (Mounk 2018). As political leaders highlight and unearth underlying cleavages until one group is the in-group versus the 'other' (McCoy et al. 2018). A pertinent case in Zimbabwe was the 2000 Referendum. The 2000 Constitutional Referendum focused on sovereignty and nationalism by emphasising the external other being whites (Onslow 2011). The referendum also proposed a change to the constitution to give more power to the president, the seizure of white-owned commercial farming land and significant restrictions on freedom of the press and electronic media (Zimbabwe Human Rights NGO Forum 2001), further threatening democratic processes (Laakso 2002, Mlambo 2013). Although the referendum was opposed. Leading up to the referendum, the government continued to use identity politics and threatened a repeat of the *Gukurahundi*, the aforementioned genocide of the Ndebele people in 1982. (Laakso 2002).

Therefore, the context in Zimbabwe provides a unique opportunity to analyse identity-based behaviours within a socio-political, ethnic, and historical framework. Although the context is unique to Zimbabwe, the relevance of such research extends beyond Zimbabwe's borders. The sentiments can be seen across the globe. The issues of ethnicity, nationalism and identity struggles can be observed in many other countries. Moreover, Zimbabwe provides an interesting research area to focus on by adding to the literature by using non-WEIRD (Western, Educated, Industrialised, Rich and Democratic) countries (Chuah et al. 2023, Henrich et al. 2010).

2.2.2 Us vs Them: Discriminatory behaviour

Social identity theory posits that cooperation and trust are more likely amongst members with the same identity (Tajfel et al. 1971). Akerlof & Kranton (2000, 2005) identified social identities as influencing economic behaviours, where agents gain positive utility from

behaviour that conforms with their group. Therefore, social identities form the basis for taste-based and statistical discrimination. According to the theory, various social identities, such as ethnicity, race, religion, nationality, gender, and artificial identities, can influence social preferences and behaviour, which has been observed in the experimental literature (Abbink & Harris 2012, Carpenter & Cardenas 2011, Chuah et al. 2013, Etang et al. 2011, Kim et al. 2017, Slonim & Guillen 2010, Chen & Li 2009). A study by Ahmed (2010) used a trust game and dictator game where the surname of the other player is revealed as either non-European or European in origin, specifically a traditionally Swedish surname or non-Swedish. Swedish men in the dictator game were found to display discriminatory behaviour toward non-Swedish players, and non-European participants did not show discriminatory behaviour towards others from a foreign background. Likewise, by revealing the surname of the other player, Der Merwe & Burns (2008) found that revealing the other player's racial identity via surnames significantly impacts the contribution size. White participants showed higher signs of in-group favouritism and out-group discrimination towards Blacks. In contrast, Black participants showed minor variation in contribution between the different surnames. In general, using this approach, researchers have identified different behaviour between those whose surnames resonate with players and those who do not.

Several studies have used photos to reveal the identity of the other player (Burns 2004, Eckel & Grossman 2004). Before accounting for racial diversity measures, Burns (2004) found that Black and Coloured participants are both less trusting of and less trustworthy towards Black senders and receivers, respectively. At the same time, White students did not engage in exchange when the other player was Black, therefore choosing to discriminate and foregoing a potential increase in endowment for interacting with Black players. Fong & Luttmer (2011) found that photos of black charity recipients did not change behaviours but changed perceptions, as participants perceived black recipients as less worthy of charity. Hence, stereotyping or statistical discrimination influences charitable giving. An alternative approach to revealing the identity of the players is through the experimenters explicitly identifying the groups being investigated (Fershtman et al. 2005, Carpenter & Cardenas 2011, Etang et al. 2011). An experiment on villagers in Cameroon, using a trust game, triple dictator Game, and risk game, found that villagers transfer significantly more to co-village members (Etang et al. 2011). In this case, participants were explicitly informed by the experimenter if the other player was someone from their village or another village for both games. The authors found that senders are more inclined to transfer their endowment when the other player is from the same village.

Beyond ethnic and racial proximity, experiments have also tested other identity groups. An experiment where half of the participants were students from a private Colombian university and the other half were students in the US in a private college Carpenter & Cardenas (2011), compared cooperative behaviour between homogeneous and intercultural players. The authors found that interculturally interactions significantly change behaviour. Experiments assessing the impact of religion on prosocial behaviour have observed mixed results, both between religious groups and between the religious and non-religious (Ahmed & Salas 2009, Anderson & Mellor 2009, Benjamin et al. 2016, Karlan 2005, Tanaka & Camerer 2016). In addition, Bigoni et al. (2019) found differences in belief perceptions between the Northern and Southern regions in Italy that impacts prosocial behaviour. Likewise, using Belgium Flemish and Walloon Universities (Fershtman et al. 2005) and University students in Malaysian, Britain and China (Chuah et al. 2007) the authors found in-group favouritism and out-group discrimination.

An important factor when considering discriminatory behaviour is the context, environment, composition of the groups and salience of the groups (Böhm et al. 2020, Bénabou & Tirole 2011*a*). For example, Chuah et al. (2007) found that Malaysian proposers playing on home soil appear to be sensitive to the nationality of their opponents. Moreover, Balliet et al. (2014) found that greater association with the in-group leads to more significant in-group favouritism. Regarding the composition of groups Burns (2004) observed that areas of high ethnic diversity and higher interactions between in-group and out-group members lead to less discriminatory behaviour in South African High Schools. Likewise, a study that tests trust and reciprocity between Belgian and Turkish participants, the largest ethnic minority in Belgium (Bouckaert & Dhaene 2004), found that trust and reciprocity are equal regardless of identity. Therefore, the frequency of interaction and the strength of identification may play a role in statistical and taste-based discrimination.

Notably, Robinson (2016) identified that if participants increasingly identify with the wider national identity, this leads to a fall in discriminatory behaviour. Similarly, Abbink & Harris (2019) compare rival and non-rival groups and find that greater association with the in-group leads to greater in-group favouritism. Although the extant literature investigates various identity groups, the prevalence, ratio, and the degree and extent these groups interact may largely impact the degree of discriminatory behaviour. Therefore, to clearly disentangle out-group discrimination, we include the respective out-group and a neutral or non-rival group (see Abbink & Harris (2019), the Zimbabwean receivers. Participants were informed that the Zimbabwean receiver had a 50% probability of being either Ndebele or Shona.

2.2.3 Group polarisation

Following social categorisation, subsequent group polarisation can arise (Mackie 1986). Social divisions may be caused by underlying polarisation. The causes and effects of polarisation have been debated in the literature. Some of this controversy arises from varying definitions and conceptualisations of polarisation. For instance, the literature on ethnic polarisation (Montalvo & Reynal-Querol 2005, Haagsma & Mouche 2013) and fractionalisation (Alesina et al. 8, Fearon 2003) focus more on linguistic differences, geographical distances, and measures of inequality. In other words, polarisation is generally used to analyse diversity between people groups. However, other scholars have identified polarisation as spreading towards the extremes of a bipolar spectrum (Duclos et al. 2003, Zhang & Kanbur 2009, Foster & Wolfson 2010, Baliga et al. 2013). Likewise, this paper focuses on the relational processes of polarisation and how people relate to one another (Lamont & Molnár 2002, Tilly 2004). *Polarisation* herein is characterised by homogeneity within groups, heterogeneity between groups, alternative groups having an insignificant effect on behaviour and greater conformity to the extremes (Mackie 1986, Sunstein 2019*a*, Esteban & Ray 1994, Iyengar & Westwood 2015, Sunstein 1999). Given the context between these two groups, we argue that these ethnicities are polarised.

The research has picked up on the different facets of polarisation, from group behaviours to social norms. For instance, experimental research by Cason & Mui (1997) focuses more on group conformity. Cason & Mui (1997) used the dictator game to identify group polarisation and whether the individuals' decisions will be significantly different from the group's decisions by varying the dictator game. The authors observed that group dynamics impact economic decision-making, teams are more altruistic than individuals. In a similar vein, LeBas (2018) and Talisse (2022) have found that polarisation in the right environment can positively impact institutions and democratic practices. However, Luhan et al. (2009) found that group environments are more selfish, and those selfish players have a greater impact on the group - moving the group to the extreme.

Similarly, Sunstein (2019*a*) found that first movers who are more selfish or polarised in discussions can pull group members towards the extreme. In addition, seminal research by Mackie (1986) observed that information from a representative of the group leads to more extreme social norms than information from a non-categorised individual. Therefore, we are also interested in investigating the degree of pluralistic ignorance. *Pluralistic ignorance* relates to the scenario when individuals' own norms and perceptions differ from group norms

(Smerdon et al. 2020, Miller & McFarland 1987, Bursztyn et al. 2020), these misperceived norms can identify polarisation as group norms appear more extreme than the individual (Mackie 1986, Sunstein 1999). This paper combines experimental findings from two-player and multi-player games with self-reported polarisation measures to identify the extent to which this polarisation has occurred and the effects on discriminatory behaviour.

Our examination of polarisation is threefold: We compare the differences between an individual's own and group perceptions to identify the movement to group extremes. We also test the degree to which a neutral or non-rival group impacts behaviours. Lastly, we use self-reported polarisation measures to further indicate the extent of affiliation - social identity salience - with the in-group as opposed to the out-group.

2.2.4 Social norms dictate behaviours

Previous research has established that an individual's behaviour across a broad spectrum of contexts is affected mainly by incentives, norms and customs. The formation of norms is ubiquitous in all societies (Elster 1989, Green 2012, Krupka & Weber 2013, Ostrom 1998). Social norms are established from connections and exchanges with others; from this, expectations regarding appropriate context-specific behaviour are developed (Ostrom 1998, Fehr & Schurtenberger 2018). As outlined by Elster (1989), social norms focus on the appropriate or inappropriate action, focusing on the action rather than the outcomes. In addition, norms are shared and recognised in a group of people. These norms are self-enforced in social contexts; hence different settings such as diverse types of individuals, locations, cultures, and time will affect the social norms adopted (Ostrom 1998). Seminal research by Krupka & Weber (2013) identified a means of measuring social norms using incentivised experiments via the social norms elicitation task. With the underlying assumption that an individual's utility increases as their material payoff increases. The amount of payoff received depends on their actions being in accordance with the social norms. We test social norms using the social norms elicitation task established by Krupka & Weber (2013); however, we use real-world vignettes to glean an understanding of norms and behaviour that are more common and realistic to participants.

2.2.5 Norm-based interventions

The prevalence of social norms in economic decision-making also implies that drawing upon certain norms can alter behaviours. Thus, social norms and subsequent framing of social norms can be practical policy tools. Shifts in social norms can be catalysts for change, improved health decisions, combating littering and further enforce institutions (Green 2012, Acemoglu & Jackson 2017, Börzel & Risse 2016, Thaler & Sunstein 2008, Kallgren et al. 2000). To date, behaviourally informed policies have mainly focused on behavioural nudges or sludges, information techniques, and defaults (Hansen & Jespersen 2013, Thaler 2018, Thaler & Sunstein 2008). *Nudges* are policies that preserve individuals' liberty whilst steering individuals to a targeted direction (Sunstein 2019*a*,*b*). In contrast, behavioural *sludges* create frictions to inhibit a particular behaviour from achieving desired policy outcomes (Thaler 2018, Lorenz-spreen et al. 2020). Behavioural policy tools, such as these, have varied in application from environmental, education and health-related issues (Abrahamse et al. 2005, Barak-Corren 2021, Miller & Prentice 2016, Allcott 2011). Behaviourally informed interventions constitute an updated methodological toolkit available for policies and initiatives. In contrast with traditional incentivised policies that may crowd out social preferences (Bowles & Polania-Reyes 2012).

Utilising norm-based intervention can anchor participants to a specific reference point (Bruns et al. 2018). The 'social' element of norms means that they should be shared across society. Norm-based interventions can take the form of either *descriptive* or *injunctive* (Cialdini et al. 1991). *Descriptive* norms prescribe typical behaviour and the likelihood of others doing the same. In contrast, *injunctive* are those norms that 'prescribe what one should do or should not do'. Nagatsu (2015) notes that social-based norms can impact social behaviour via expectations or framing. Expectations create the assumption that others will behave accordingly and cooperate among conditional cooperators. Whereas framing shifts actions from 'I' to 'We', making agents consider the group's utility above or equal to their own. Therefore, referring or framing around a reference group can improve the effectiveness of norm-based nudges (Andor et al. 2022). A further explanation for why norm nudges can be effective in shifting behaviours is due to the individuals' underlying preferences as either collectivists or individualists (Mol et al. 2020, Guala et al. 2013).

A study conducted in Zimbabwe and Swaziland, Bicchieri (2015) found that norm-based interventions can alter individuals' behaviour when considering gender roles. Similarly, Bursztyn et al. (2020) found that correcting the misperceived norms regarding women in work

increased male acceptance and female participation in the labour force in Saudi Arabia. Accordingly, policymakers should not underestimate the power of social norms and the desire to fit in or our willingness to do something depending on an awareness of what others do (Bursztyn & Jensen 2017, Börzel & Risse 2016, Bursztyn et al. 2020). In this paper, we focus on injunctive norms - what others find appropriate behaviour (Cialdini et al. 1991, Farrow et al. 2017). By providing individuals with information on the norms of their group, *Heal*, a prescriptive social norm that describes what people approve of doing, and *Divide*, a proscriptive social norm which describes what people disapprove of doing. Individuals can deliberate – that is, to use system two thinking[2] – as they notice that their beliefs are either shared or not with others within their network. Subjects then ensure that they behave accordingly to comply with the norm (Bicchieri 2015, Miller & Prentice 2016).

2.3 Theoretical Framework

Group polarisation is closely linked to social identity formation and intergroup conflict. Various theories have been identified to explain in-group favouritism and out-group discrimination, such as bounded generalised reciprocity (Yamagishi & Kiyonari 2000) and the evolutionary theory of group behaviours (Gintis et al. 2003). In this paper, we focus on social identity theory (SIT). SIT posits that people either maximise in-group outcomes at the expense of out-group outcomes, even in minimal or artificial groups (Tajfel et al. 1971). Agents identify themselves based on mere self-categorisation into different groups separate from their personal or individual traits. This categorisation has an emotional response which leads to divergent behaviours between groups (Balliet et al. 2014). Likewise, Akerlof & Kranton (2000, 2005) stress the importance of identity in influencing economic behaviour. The authors suppose that economic agents gain utility from behaving appropriately with respect to their social group. Identity with a specific group will impact behaviour with in-group and out-group members (Chmura et al. 2005). This identity is internal to agents, and there may be the threat of sanction if one deviates from the group's norms; even when one is by themselves, the identity that they prescribe themselves and the associated social norms guides them to a particular behaviour, even at the expense of individual interest (Elster 1989). Hence, to avoid sanctioning and maintain positive self-esteem, agents conform to the

[2]Dual process model has shown two distinct systems. System 1 is automatic and unconscious. Whereas System 2 is more deliberative and considered, which is likely to have more prolonged behavioural changes (Dolan et al. 2012, Ölander & Thøgersen 2014, Rangel et al. 2008, Ohtomo & Hirose 2007). See John et al. (2009) for a review of systems thinking and nudges.

social identity, even at the cost of economic efficiency, often viewing their group as morally right, whilst the out-group is the opposite.

We use a simple model of utility maximisation that follows Mill & Morgan (2021), as well as the role of norm-compliance and identity mirroring Krupka & Weber (2013)[3]. Akerlof & Kranton (2000) developed an economic theory that hypothesised that different identities would alter the payoffs of certain actions. An individual $i's$ utility $U(a_i)$ depends on both their actions and the degree to which their actions comply with the prevailing norms $N(a_i)$:

$$U(a_i) = V(\pi(a_i)) + \gamma N(a_i) + (\eta)(\pi_i(a_i) - \pi_j(a_j)) \tag{2.1}$$

The component V represents the value individuals place on the material payoff of action a. The function increases in $\gamma(a_i)$, $\gamma \geq 0$, as it increases, an individual will get higher utility from complying with the norm - in other words, doing the socially appropriate action. With $N(a_i)$ being a norm function. If people were completely concerned about economic efficiency $\gamma = 0$, they would always choose the option that maximised payoff. Deviation from the social norm yields disutility. This fear of deviation or disapproval of others can generate a "psychological cost" (Fehr & Schurtenberger 2018). The "social multiplier" (Bénabou & Tirole 2011*b*, p.6), is likely to drive a shift in behaviour in line with the social norm, which we exogenously manipulate in the *Heal* and *Divide treatments*. Participants with a higher γ will be more inclined to comply with cooperative social norms even in one-shot games. However, lower levels of γ, suggest that these participants are on the threshold and are more likely to deviate from the social norm.

The ethnic identity is given by η. Let $\eta = 0$ if participants share the same ethnic identity as their opponent, and $\eta = 1$, otherwise. The utility of individual i is increasing in π_i. The decreasing utility in $\pi_j(a)$ is larger if the participants do not share the same ethnic identity as their opponent. Thus, we hypothesise that polarisation will be more significant if they are paired with an opposing ethnicity compared to a coinciding ethnicity. Therefore, agents yield higher utility from maintaining positive self-esteem by conforming to group behaviour. Polarisation leads to stronger identification with the in-group, which will impact discriminatory behaviour. See for example, Appendix A, for additional theoretical framework.

Hypothesis 1 *Participants discriminate more if they are paired with an opposing ethnic group member.*

[3]For all four experimental games we maintain the same definitions

Hypothesis 2 *The more strongly individuals identify with their ethnicity (or the more polarised), the greater the discriminatory behaviour.*

Hypothesis 3 *Following the* Heal *treatment, participants will be more prosocial towards out-group members.*

Hypothesis 4 *Following the* Divide *treatment, participants will be less prosocial toward out-group members.*

2.4 Experimental Method

In the following, we present the experimental design, the procedure and other measures of interest. We use lab-in-the-field experiments conducted in Zimbabwe. To ensure the robustness and external validity of our research we use several experimental games to disentangle willingness to coordinate, cooperate, exhibit altruism and trust. In line with the literature, we use several experimental games, beliefs and social norm (see for example, (Bigoni et al. 2019, Etang et al. 2011, Kim et al. 2017, Kimbrough & Vostroknutov 2016) to capture pro-social behaviours. Moreover, Wang & Navarro-Martinez (2023) find that aggregating pro-social behaviour measures can enhance the external validity of experimental research. Therefore, when investigating social norms behaviours, we use an aggregated measure of behaviour in each experimental game, respectively.

2.4.1 Prisoner's Dilemma (PD)

The prisoner's dilemma is a two-player game where the players can choose between two strategies: cooperate (C) or defect (D). Each participant knows how much they could gain, but this depends on the other participant's choice (see Table 3.1). Decisions are made simultaneously. Thus an individual is not aware of the decision made by the other individual until after they both make a decision. Participants who choose an outcome with the least difference between themselves and the other player indicate a preference to cooperate (C, C); this is the Pareto-dominant strategy since there is no other option that gives both players a higher payoff. However, preferences to defect are indicated by participants choosing the outcome with the highest individual payoff, even if there is a significant difference between

payoffs between participants (Cardenas & Carpenter 2008). Defecting is the dominant strategy (D, D) (see Figure A.1 for the table shown to participants). The strategy is dominant as it is the best outcome for the individual player regardless of what the other player does. The Pareto-dominant outcome (C, C) is superior to the dominant strategy as players are better off relative to the Nash equilibrium. The assumption of dominance in the prisoner's dilemma means that utility maximising agents will choose the strictly dominant outcome (Camerer 2011). Participants are tasked with making incentivized decisions between an in-group, out-group and mixed-group receiver; this indicates if preferences toward cooperation change depending on the social identity of the partner[4].

Table 2.1: Payoff Matrix (Prisoner's Dilemma)

		Player 2	
		Defect	*Cooperate*
Player 1	*Defect*	(1,1)	(10,0)
	Cooperate	(0,10)	(5,5)

The one-shot Prisoner's Dilemma measures taste-based discrimination, which is indicated by the difference in cooperation between in-group and out-group matching. We use the prisoner's dilemma because we can specifically identify taste-based discrimination. We also elicited participants' beliefs regarding the decisions of others, which measures statistical discrimination.

You Interact	Other Interacts	You Win	Opponent Wins
		5	5
		0	10
		1	1
		10	0

Figure 2.1: Prisoner's Dilemma in the experiment (in points)

[4]1 point correlates to $\$US0.50$.

2.4.2 Dictator Game (DG)

The dictator game by Roth & Erev (1995) is a decision-making task that tests if agents are purely self-interested. Player A (the dictator) is endowed with 10 points for each receiver; they can decide to keep it all or to share this endowment with Player B. If Player A allocates an amount to Player B, Player A gets to keep the remaining amount. A fair outcome would be to share the endowment equally between the two participants. In this paper, participants are tasked with making incentivized decisions between in-group, out-group and mixed-group receivers. Differences in the amount sent to the receiver who are in-group, out-group and mixed-group members indicate discriminatory behaviour - precisely, the sender's behaviour measures taste-based discrimination (Grosskopf & Pearce 2017).

Figure 2.2: Dictator Game - Sender

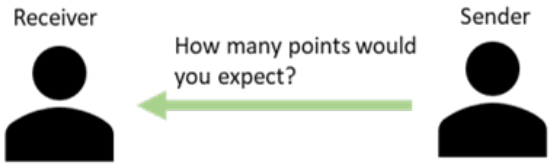


Figure 2.3: Dictator Game - Beliefs

2.4.3 Trust Game (TG)

The trust game by Berg et al. (1995) identifies the degree of trust and reciprocity between participants. The trust game is a two-player game where Player A is endowed with 10 points for each receiver. Player A is tasked with deciding how much endowment (x), if any, to send to Player B. Any amount Player A does not send, they can keep. The amount they send to

player two is multiplied by 1.5. Player B receives the amount sent by Player A ($3x$). Player B decides how much to return (y) to Player A, with $y \leq 3x$.

Final payoffs are $P_A(x,y) = E - x + y$ for player A, and $P_B(x,y) = 3x - y$ for player B.

The amount transferred (x) measures player A's trust. Any positive contribution from Player A indicates that they 'trust' that either the in-group, out-group or mixed-group reciprocate any positive contribution. However, Player A is vulnerable as Player B may choose to keep all positive contributions for themselves. Participants are asked to make decisions as both Player A and Player B.

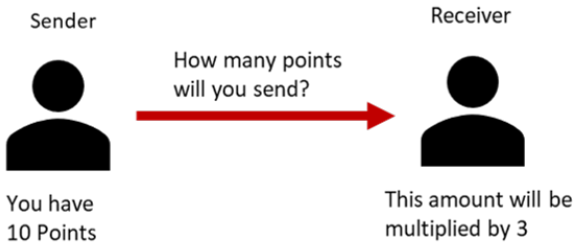


Figure 2.4: Trust Game - Sender

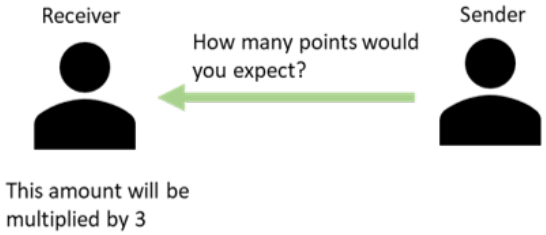


Figure 2.5: Trust Game - Beliefs

2.4.4 Public Goods Game (PGG)

Participants play the game with three other players in the public goods game (PGG) by Gächter & Fehr (1999). Participants make contributions to the common pot, the total contribution is multiplied by a factor α $(1 < \alpha < n)$, where n denotes the number of participants. The contributions will be equally re-distributed between all four participants: the PGG tests cooperativeness and the degree of prosocial behaviour.

Each participant is endowed with 10 points for each receiver and is asked to decide how much to keep or contribute to a public good. The sum of the group's payoff is multiplied by 1.5, which is shared equally between the four participants. The PGG occurred with homogeneous groups or politically mixed groups. Participants are tasked with making incentivized decisions between in-group $(n-1)$, out-group $(n-1)$ and mixed-group $(n-1)$ players. The final payoff for the game for each participant is given by:

$$\pi_i = 10 - g_i + \frac{1.5}{4} \sum_{j=1}^{4} g_j \tag{2.2}$$

The public goods game has two equilibria, a Nash equilibrium and a Pareto optimal outcome. The Pareto (or socially) optimal equilibrium occurs when you cannot make any participants better off through the reallocation of payoffs. In the PGG, this occurs when there is full cooperation, put differently, when all players contribute their entire endowment to the common pot. The Pareto optimal outcome suffers from being a weak equilibrium. The alternative non-cooperative equilibrium is stronger. If someone from the group deviates from the socially optimal outcome due to the presence of free-riders, then more and more participants will free-ride over time. At the alternative equilibrium, if an individual deviates and begins to contribute more than zero, contributions will quickly go back to zero. Therefore, the dominant strategy is to contribute zero of their points to the common pot. We use the public goods game as it examines taste-based discrimination in the one-shot game.

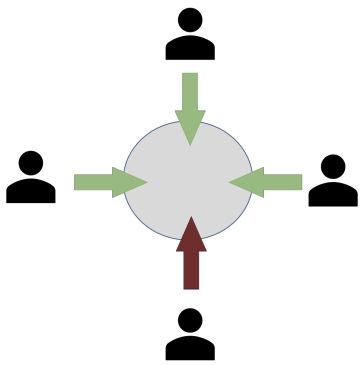


Figure 2.6: Public Goods Game - Sender

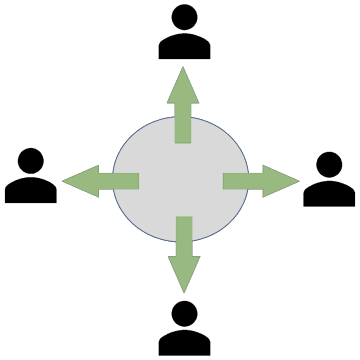


Figure 2.7: Public Goods Game - Beliefs

2.4.5 Belief elicitation

Given this paper's focus, games categorised as social dilemmas were used to identify discriminatory behaviour. We also elicited participants' beliefs regarding others' decisions for all social dilemma games. Hence, participants made decisions as senders and were also tasked with considering the receiver's decisions. Belief elicitation occurs after the one-shot games listed above to identify an agent's belief about the state of the world. Participants must consider how much of the endowment the receiver would send to them if their roles were reserved. This reveals their beliefs about the other players. In particular, we elicited their beliefs about in-group, out-group and mixed-group players as participants may expect less cooperative and prosocial behaviour from in-group members (Ferraro & Cummings 2007). The literature has shown that beliefs about others' behaviour can affect instances of cooperation and prosocial behaviour (Fischbacher & Gächter 2010, Gächter & Renner 2010, Guala et al. 2013). Belief elicitation was incentivizied, we matched their beliefs ex post (De Geest & Stranlund 2019, Schotter & Trevino 2014, Gächter & Renner 2010). In general, people form beliefs about others to explain divisions between groups. Therefore, belief elicitation allowed us to identify statistical discrimination.

2.4.6 Social Norms Elicitation Task

We also elicited social norms using the seminal method by Krupka & Weber (2013). The method applied deviates from Krupka & Weber (2013) by using real-world scenarios rather than the allocator game. This removes the abstract nature of game-based situations to enable more realistic ideas of social norms. Additionally, participants are asked to evaluate

each scenario three times, first considering their perceptions of the social appropriateness of the provided action, then what most people that belong to their in-group and what most people that belong to a different identity group consider socially (in)appropriate.

Participants' beliefs about social norms are incentivized, as they are made aware that their decisions will be matched with the modal response of Shona and Ndebele participants. The incentivization of the social norms task is vital as norms are perceived collective behaviours (Akerlof & Kranton 2005, Elster 1989, Green 2012, Ostrom 1998). Nonetheless, in this study, participants are asked to evaluate their perceptions before that of the group in which they belong and the out-group. All participants are asked to evaluate the scenarios in this way to provide insight into the difference between personal convictions and social norms [5]. We used a within-subject design to control for idiosyncrasies in participants (D' Adda et al. 2016). Participants evaluated the following scenarios:

- *Pay Scenario - A manager of a company shows favouritism to certain employees. The manager gives a pay rise to people who are the same ethnicity as the manager.*

- *Tax Scenario - Suppose that a shop owner falsely reports sales, which leads to them paying less tax.*

- *Family Scenario - Suppose someone in your family wants to get married to someone in the opposite ethnicity.*

- *Briber Scenario 1 - You see a manager bribing a public official to get a contract. How appropriate is it to report the manager.*

- *Bribe Scenario 2 - A citizen is made aware of this bribe. How appropriate is it to punish the manager.*

[5]Note that the social norm elicitation task occurred after the experimental games. However, D' Adda et al. (2016) found limited evidence of order effects of norm elicitation task.

2.4.7 Stage 2

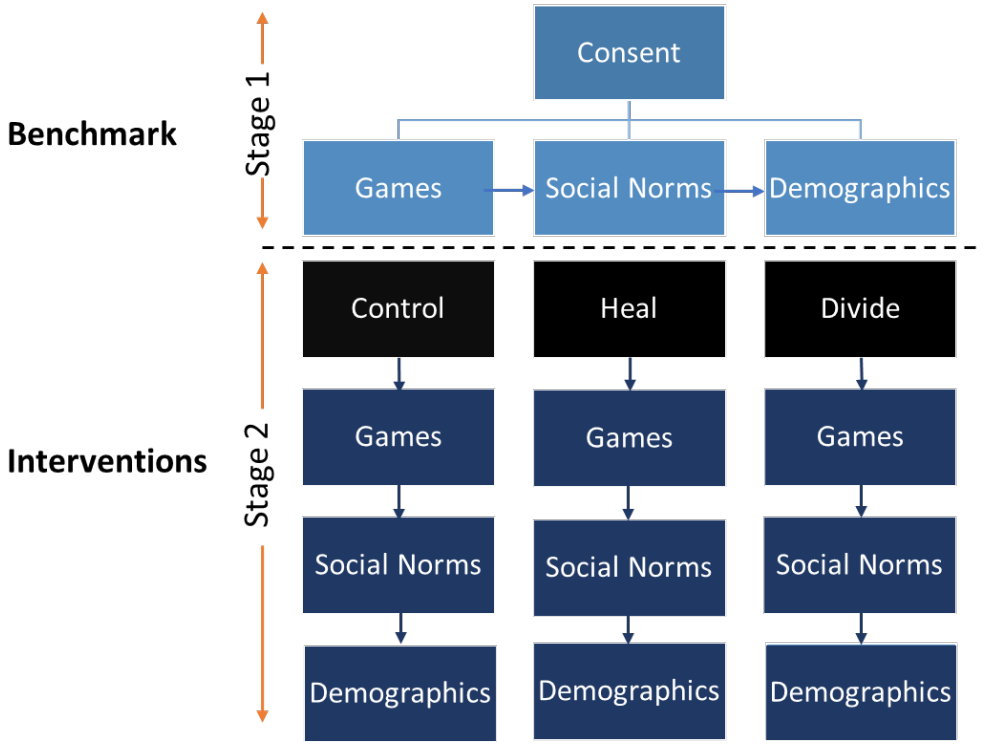


Figure 2.8: Experimental Summary

In the second part of the experiment, participants are randomly assigned to the treatment conditions. Participants receive information about the prevailing social norms. Therefore, we can explicitly identify the direction of the effects of the norm-based intervention. More specifically, if it's any information or what information that matters. Participants in the first treatment group are given feedback on previous respondents with shared social norms from the in-group and out-group, labelled the *HEAL* treatment. In the second treatment group, participants are given feedback on different and diverging social norms between the in-group and out-group; this treatment is labelled the *DIVIDE* treatment. Norms surrounding marital linkages between Shona and Ndebele people were used as Muzondidya & Ndlovu-Gatsheni (2007) found that marriages were used as an instrument to either facilitate or further fracture connections in Zimbabwe, see Figure 2.9 and Figure 2.10. The *Control* group receives no information regarding social norms.

After being placed into one of the three treatment groups, participants are given the same games, social norms elicitation task and post-experiment questionnaire as those in part one. See Figure 2.8 for a summary of experiments. The random allocation of the intervention

occurs before the games are played so that social norms are primed. This brings norms to the forefront and increases norm compliance. Besides the different information shown to participants, the remainder of the experiment is unchanged across treatments.

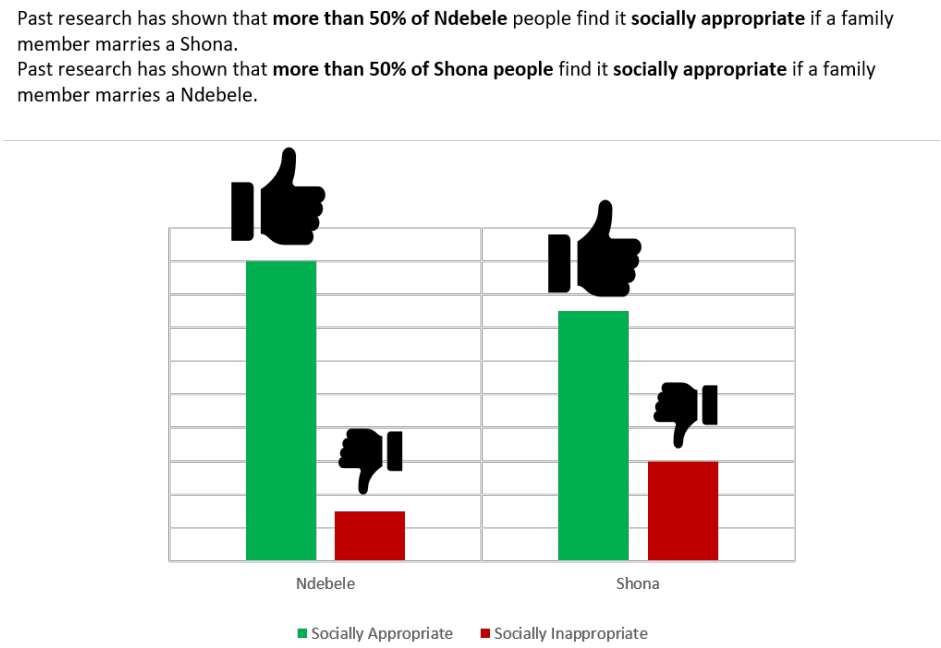


Figure 2.9: Social Norm nudge - HEAL

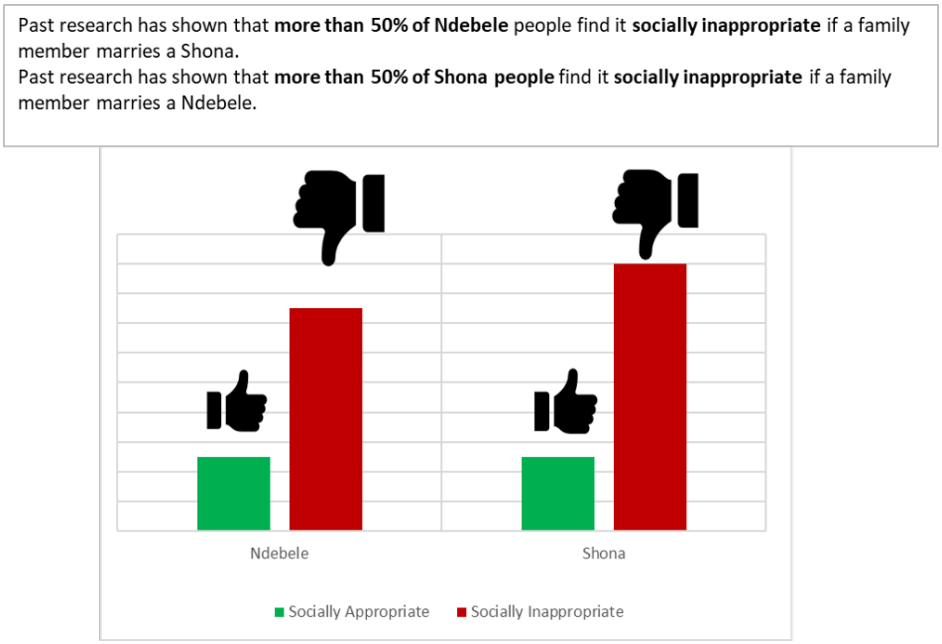


Figure 2.10: Social Norm nudge - DIVIDE

2.4.8 Procedure

The experiments were conducted in Zimbabwe, over several sessions at Marondera Agricultural University, Harare Polytechnic, Nyadire Teaching College, Lupane State University, and University of Zimbabwe [6]. At the beginning of the in-person sessions, participants received written instructions and were given the opportunity to ask questions privately [7]. They were informed that they would receive a fixed participation fee ($US2) and an additional payoff that would differ for each participant. Participants were truthfully told that one of their choices would be randomly selected at the end of the experiment to determine their payoff. See Appendix A for instructions. Participants joined the experiment voluntarily by responding to an email and flyer sent to all students prior to the experimental sessions, alternatively, after being invited through Prolific. All participants were tasked with undertaking the prisoner's dilemma, dictator game, trust game, and public goods game [8] and the social norms elicitation task. Participants did not receive any feedback from the games undertaken.

Following these tasks, all participants completed a post-experiment questionnaire. In the post-experiment questionnaire, we asked participants questions about their perceptions to shed light on their subjective views on their identities. Firstly, participants were asked if they were Ndebele or Shona. Secondly, participants were given the option to decide if they identify as either *African*, *Ndebele*, *Shona* or *Zimbabwean*. This allows us to glean the extent of affiliation of the identity groups. For instance, participants may prefer the social categorisation of being Zimbabwean or African. They then assess to what extent they align with their in-group identity. Thirdly, we asked participants to evaluate how proud they are to be their in-group identity; the question is coded to run from 1 to 4, 1 being "not proud at all" to 4 being "very proud". Additionally, to ensure that the heterogeneity between these two groups of interest is not neglected, data such as sex, education, ethnicity, employment status, household size, and marital status, as well as information such as age and income variables, are collected to use as controls in the analysis of the research.

[6] Experiments were conducted in English. However, Ndebele and Shona translations were available for participants who needed them.

[7] The instructions used were neutral and aimed not to detail the objectives of the experiment.

[8] The games are randomised to prevent any ordering effects

2.5 Results

The following analysis begins by testing the differences in behaviours between the two identities. We also consider the role of self-reported identity preferences for cooperation as a proxy for polarisation. In our analysis, we include these in our main regression models. Following this, the effect of the social norm-based intervention is analysed. Lastly, we evaluate the influence of the social norm-based intervention on behaviours and norm evaluations.

Table D1 summarises the baseline characteristics of the participants in the experiment. Consistent with the regional context, we have more Shona participants than Ndebele. Over half of the participants are Shona (Comparative Constitution Project 2018). Most participants across all treatments are between 18 and 44 years old. Similarly, the majority of participants are in or have a University Degree. [9]. We used a student population to ensure a sufficient level of literacy. Students are commonly used in experimental research. As noted by Belot et al. (2015) and Snowberg & Yariv (2021), a comparison of the experimental literature reveals little difference between student and non-student populations. Similarly, Lane (2016), found that both students and non-students seem to exhibit comparable levels of discrimination. In addition, the demographic characteristics of the students in our study revealed a greater level of non-traditional-aged students. The inclusion of a considerable number of mature students provides our student sample with greater variation.

[9] A note of caution is due here. Since the data collected was halted due to the onset of the COVID-19 pandemic, we were unable to conduct the experiments with the anticipated number of participants. This also impacted the distribution of characteristics across treatments.

Table 2.2: Summary Statistics

	All	**Control**	**Heal**	**Divide**
Identity				
Ndebele	20.58	23.66	10.20	11.76
Shona	79.42	76.34	89.80	88.24
Self-identity (%)				
African	44.19	48.90	32.65	29.41
Ndebele	11.94	6.17	8.16	26.47
Shona	4.84	3.08	8.16	11.76
Zimbabwean	39.03	41.85	30.61	32.35
Socio-economic characteristics				
Age (%)				
18-24	45.67	57.20	17.31	13.16
25-34	27.16	27.43	28.85	31.58
35-44	10.45	26.34	15.38	25.79
Female (%)	62.28	62.95	67.35	50.00
Single (%)	62.79	71.05	44.90	32.35
Household size	4.59	4.81	4.14	3.79
Risk	7.30	7.50	6.73	6.79
Tertiary education (%)	88.10	88.16	89.80	85.29
Employment status (%)				
Student	58.33	69.87	24.49	29.41
Employed part-time	6.09	4.80	6.12	14.71
Employed full time	24.04	15.28	51.02	44.12
Self employed	4.49	3.93	33.3	2.94
Polarisation measures				
Aligned with Ndebele identity (%)	39.68	48.46	12.24	20.59
Aligned with Shona identity (%)	76.77	77.53	75.51	73.53
Proud of identity	86.45	92.95	65.30	73.53
Observations	423	270	78	75

Note: Baseline summary statistics of participant characteristics in the experiment.

We begin by testing the differences in the self-reported polarisation measures between the two ethnolinguistic identities. There is a statistically significant difference in reported alignment with the Ndebele identity between Shona and Ndebele participants (Wilcoxon rank-sum test p=0.000). Similarly, individuals who align highly to the Shona identity (Wilcoxon rank-sum test p=0.000). However, we find no statistically significant differences in being proud of

being Ndebele or Shona (Wilcoxon rank-sum test p=0.227). Hence, within the sample of participants, there is a clear difference between those who identify as Ndebele and those who identify as Shona. Nevertheless, being proud of that identity has similar evaluations between the two ethnicities.

Throughout the analysis, discriminatory behaviour is identified as having lower contributions. In contrast, favouritism is identified as greater contributions. We anticipate that this will be motivated by having a shared identity. We now turn to the levels of discriminatory behaviour observed in each economic game.

2.5.1 Prisoner's Dilemma

In the prisoner's dilemma, participants chose between cooperating or defecting, labelled as one if the participant decided to coordinate, zero otherwise. We find that the beliefs about Zimbabwean and Shona participants' likelihood of coordination are 0.60 and 0.56, respectively, which is less than actual coordination with Zimbabwean and Shona receivers which is 0.72 and 0.72, respectively. Whereas beliefs about Ndebele participants' likelihood of coordination is 0.48, which is greater than actual coordination with Ndebele players 0.40, see Figure 2.11. Compared to the mean level of coordination, Shona participants coordinate more than the mean when the receiver is Zimbabwean and Shona, respectively, but less than the mean level of coordination for Ndebele participants, see Figure 2.12. In comparison, Ndebele participants contribute more than the mean level of coordination when coordinating with Ndebele receivers (in-group) only.

We use Wilcoxon rank-sum tests to test differences in coordination in the prisoner's dilemma game. Coordination is statistically different between the two ethnic groups when coordinating with a Ndebele participant (p = 0.000). There is a statistically different level of coordination between the two ethnicities when coordinating with the Shona receiver (p=0.007). There are no significant differences between the two ethnic groups when coordinating with a Zimbabwean receiver (p= 0.292). Therefore, we find evidence of polarisation and taste-based discrimination because we observe differences in contributions between the two ethnicities depending on the receiver's identity. Consistent with our hypothesis, we find no statistical difference between the two groups if the receiver is Zimbabwean.

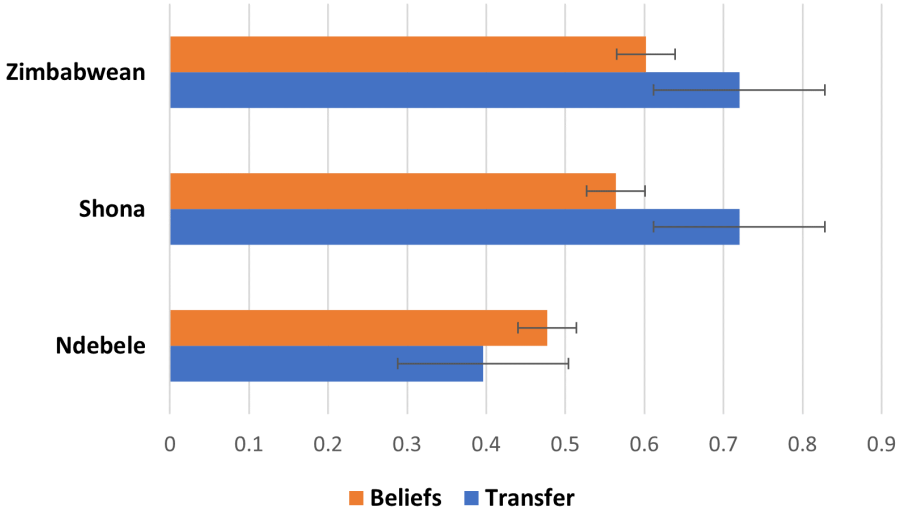


Figure 2.11: Prisoner's Dilemma (Mean coordination and beliefs)

Note: The figure shows the percentage of participants choosing cooperation (=1) and the beliefs of others choosing to cooperate. Orange bars present beliefs and the blue bars present the amount sent for each receiver (Zimbabwean, Shona and Ndebele). Error bars indicate 95% confidence intervals.

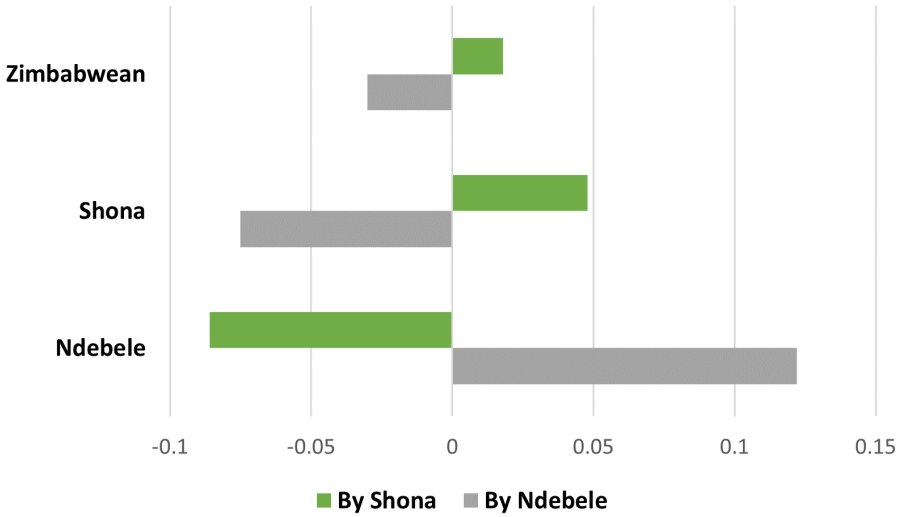


Figure 2.12: Prisoner's Dilemma (mean-centred)

Note: Zero line shows the pooled mean coordination as a benchmark. Orange bars show differences from the mean likelihood of coordination by Shona participants. Blue bars show differences from the mean likelihood of coordination by Ndebele participants.

In Table 2.3, we analyse coordination behaviour in the prisoner's dilemma and identity. In columns (1) and (2), the dependent variable is the likelihood of coordination with Ndebele receivers. In columns (3) and (4) the dependent variable is the likelihood if coordination with Shona receivers. Lastly, in columns (5) and (6), the dependent variable is the likelihood

of coordination with Zimbabwean receivers. *Shona* denotes a dummy variable with a value of one if the participant is Shona and zero if the participant is Ndebele. We find evidence of in-group favouritism and out-group discrimination. However, our findings are not robust to the inclusion of controls. We find that beliefs about Ndebele players significantly influence behaviours.

We obtain consistently insignificant results from our identity salient measures[10]. We do, however, observe that individuals highly aligned with the Ndebele identity are more likely to coordinate with their in-group (Ndebele receivers) and less likely with the out-group (Shona receivers). In addition, individuals who are proud of their identity (*proud*) are more likely to coordinate with Zimbabwean receivers.

Table 2.3: Probit regression Prisoner's Dilemma by contribution

Dep. variable: Coordinate(=1, 0 otherwise)	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(3)	(4)	(5)	(6)
Shona(=1, Ndebele=0)	-0.320***	-0.063	0.554***	0.083	0.044	-0.092
	(0.103)	(0.150)	(0.214)	(0.236)	(0.083)	(0.200)
Beliefs	0.419***	0.440***	0.019	0.041	0.294	0.271
	(0.125)	(0.107)	(0.168)	(0.143)	(0.187)	(0.199)
Self-identity (Base = African)						
Ndebele		0.380		0.737		0.000
		(0.383)		(0.501)		(.)
Shona		-0.409		-0.131		-0.075
		(0.435)		(0.639)		(0.355)
Zimbabwean		0.146		0.022		-0.012
		(0.231)		(0.319)		(0.212)
Polarisation						
Shona align		-0.009		0.149		-0.031
		(0.090)		(0.099)		(0.088)

Table 2.3 – continued from previous page

	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(3)	(4)	(5)	(6)

[10]We are unable to report coefficients for *Ndebele* self-identity due to insufficient observations.

Ndebele align		0.136**		-0.259*		-0.126
		(0.065)		(0.151)		(0.121)
Proud		0.199		0.289		0.383*
		(0.197)		(0.321)		(0.214)
Controls	NO	YES	NO	YES	NO	YES
Constant	0.057	-1.790	-0.363	-1.337	0.355**	-0.739*
	(0.219)	(1.202)	(0.329)	(0.914)	(0.172)	(0.443)
N	278	266	278	266	278	254

Note: Standard errors clustered at the subject level in parentheses.*** p<0.01,** p<0.05,* p<0.10.

2.5.2 Dictator Game

Figure 2.13 shows the share of dictators that choose to cooperate and their corresponding beliefs. Of their endowment of 10 points, for each receiver, beliefs are greater than actual transfers. Beliefs for Zimbabwean, Shona and Ndebele receivers are 6.25, 5.66 and 5.15, respectively. Transfers for Zimbabwean, Shona and Ndebele receivers are 5.84, 4.68 and 5.19. In Figure 2.14, zero indicates the mean contributions for all receivers. Compared to mean contributions, Shona participants contribute less to Zimbabwean and Ndebele receivers. However, Shona participants contribute more than the average contributions to the in-group. Whereas Ndebele participants contribute more than the mean level of contributions to Zimbabwean (non-rival group) and Ndebele (in-group) receivers, but less for Shona receivers (out-group).

Testing for differences in contributions by ethnic identity. We find a statistically different contribution between the two ethnicity groups when contributing to Ndebele participants (Wilcoxon rank-sum test p=0.000). However, we find no statistically significant differences in contributions when contributing to Shona participants and Zimbabwean participants by ethnic groups (Wilcoxon rank-sum test p=0.443 and p=0.338, respectively).

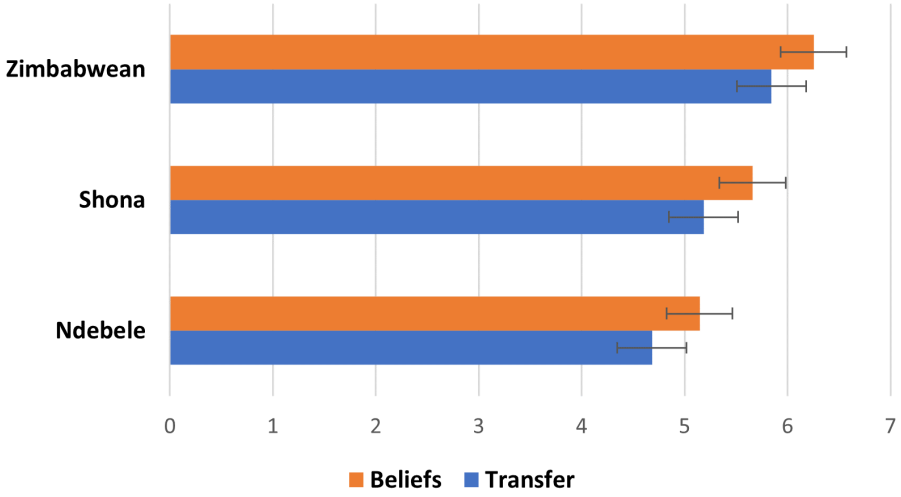


Figure 2.13: Dictator Game (Mean contributions and beliefs)

Note: The figure shows the amount the dictators (senders) choosing to cooperate with and the beliefs about the corresponding receiver. Orange bars show beliefs, and the blue bars present the amount sent for each receiver (Zimbabwean, Shona and Ndebele). Error bars indicate 95% confidence intervals.

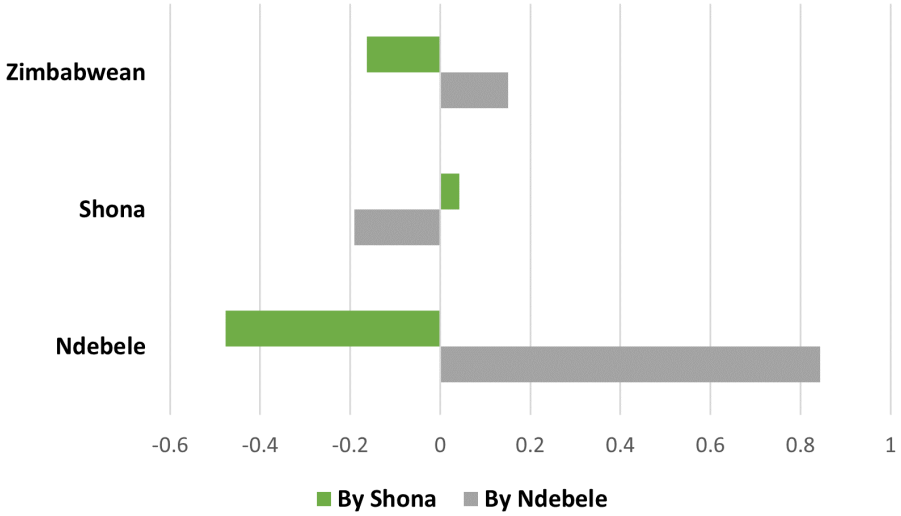


Figure 2.14: Dictator Game (mean-centred)

Note: Zero shows the mean as the benchmark. Orange bars shows differences from mean transfers by Shona participants. Blue bars show transfers different from the mean by Ndebele participants.

Table 2.4 presents OLS model estimates of cooperation on identity and the intervention, including additional controls, polarisation, sex, employment status, age, household size, marital status and risk. In columns (1) and (2), the dependent variable is the amount of cooperation (or transfers) with Ndebele receivers. In columns (3) and (4), the dependent variable

is the amount of cooperation (or transfers) with Shona receivers. Lastly, in columns (5) and (6), the dependent variable is the amount of cooperation (or transfers) with Zimbabwean receivers. We find evidence of discriminatory behaviour, Shona dictators (participants) transfer less than Ndebele dictators when the transfer goes to Ndebele receivers. Conversely, Shona dictators favour their in-group. In addition, we find that Shona participants transfer more than Ndebele dictators to Zimbabwean receivers in column (6). Beliefs regarding others' behaviours significantly impact transfers in the dictator game. Beliefs regarding each respective receiver lead to greater amounts sent to the receiver.

Using self-reported identity measures, we find that participants who self-identify as Ndebele, when given a broader choice of identities, significantly reduce transfers to Zimbabwean receivers than individuals who chose to identify as African. Whilst the other polarisation measures do not significantly impact trust.

Table 2.4: OLS regression Dictator Game by contribution

Dep. variable: Amount sent to receiver	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(3)	(4)	(5)	(6)
Shona(=1, Ndebele=0)	-1.046***	-0.620*	0.183	0.530*	-0.175	0.497**
	(0.190)	(0.298)	(0.277)	(0.249)	(0.215)	(0.148)
Beliefs	0.480***	0.463***	0.520***	0.536***	0.612***	0.632***
	(0.041)	(0.035)	(0.033)	(0.036)	(0.051)	(0.064)
Self-identity (Base = African)						
Ndebele		0.234		0.100		-0.760**
		(0.586)		(0.461)		(0.315)
Shona		0.177		-0.090		-0.169
		(0.583)		(0.630)		(0.441)
Ndebele		0.341		-0.029		-0.371
		(0.307)		(0.221)		(0.324)
Polarisation						
Shona align		-0.088		-0.283		-0.282

Table 2.4 – continued from previous page

	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(3)	(4)	(5)	(6)

		(0.157)		(0.228)		(0.175)
Ndebele align		0.157		-0.053		0.052
		(0.126)		(0.153)		(0.110)
Proud		-0.041		0.410		-0.005
		(0.124)		(0.229)		(0.208)
Controls	NO	YES	NO	YES	NO	YES
Constant	4.054***	2.414*	1.911***	-0.016	2.184***	0.582
	(0.458)	(1.136)	(0.566)	(1.117)	(0.556)	(1.497)
N	344	303	344	303	344	303

Note: Standard errors clustered at the subject level in parentheses. ***p<0.01,** p<0.05,* p<0.10.

2.5.3 Trust Game

Figure 2.15 shows beliefs and actual transfers for each receiver. Participants' beliefs of other Zimbabwean, Shona and Ndebele receivers are 6.23, 5.84 and 5.44, respectively. Actual transfers for Zimbabwean, Shona and Ndebele receivers are lower than beliefs at 6.19, 5.49, and 5.16, respectively. In Figure 2.16, zero indicates the mean contributions for all receivers. Compared to mean contributions, Shona participants contribute less to Zimbabwean and Ndebele receivers. However, Shona participants contribute more than the average contributions to the in-group. Whereas Ndebele participants contribute more than the mean level of contributions to Zimbabwean (mixed group) and Ndebele (in-group) receivers, but less for Shona receivers (out-group).

We also examine differences in contributions by ethnic identity using a non-parametric rank-sum test. In the trust game, we find a statistically different contribution between the two ethnicity groups when contributing to Ndebele participants (p=0.000). However, we find no statistically significant differences in contributions when contributing to Shona participants and Zimbabwean participants by ethnic groups (p=0.206 and p=0.135, respectively). Therefore, we find evidence of Taste-based discrimination when the receiver is Ndebele.

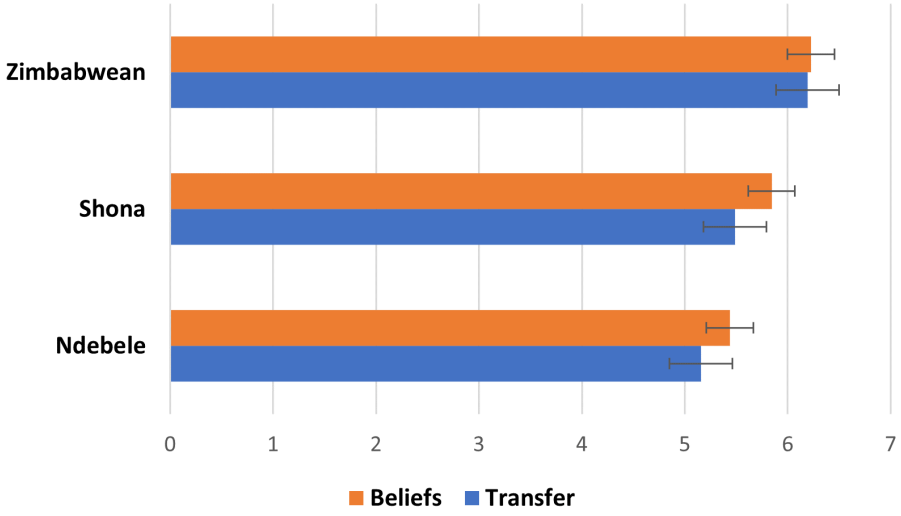


Figure 2.15: Trust Game (Mean contributions and beliefs)

Note: The figure shows the amount the sender chooses to trust (transfer) and the beliefs of the corresponding receiver. The orange bars show beliefs, and the blue bars present the amount sent for each receiver (Zimbabwean, Shona and Ndebele). Error bars indicate 95% confidence intervals.

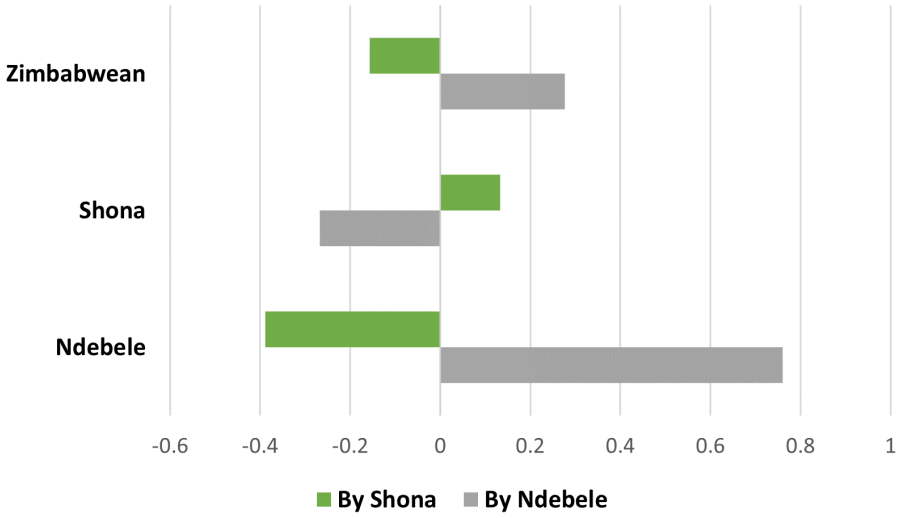


Figure 2.16: Trust Game (mean-centred)

Note: Zero shows the mean as the benchmark. Orange bars show differences from mean transfers by Shona participants. The blue bars show transfers different from the mean by Ndebele participants.

Table 2.5 presents OLS model estimates of cooperation on identity and the intervention, including additional controls, polarisation, sex, employment status, age, household size, marital status and risk. In columns (1) and (2), the dependent variable is the amount of cooperation (or transfers) with Ndebele receivers. In columns (3) and (4), the dependent variable is the

amount of cooperation (or transfers) with Shona receivers. Lastly, in columns (5) and (6), the dependent variable is the amount of cooperation (or transfers) with Zimbabwean receivers. Shona trustors (participants) transfer more to other Shona receivers. However, this result is not robust to the inclusion of controls; see columns (3) and (4). We also observe that Shona trustors transfer less than Ndebele receivers to Zimbabwean receivers. We observe that beliefs significantly determine behaviours. Beliefs about Ndebele, Shona and Zimbabwean players in the role of sender increase transfers to Ndebele, Shona and Zimbabwean receivers.

Surprisingly, individuals who self-identify as Ndebele transfer more to Shona receivers. Whereas individuals who self-identify as Shona contribute less to Ndebele receivers. We speculate that this finding is due to the potential for reciprocity, with individuals choosing the Ndebele identity from a selection of options, anticipating that Shona participants are more likely to reciprocate. Regarding the self-reported polarisation measures, being proud of one's identity increases trust toward Shona receivers.

Table 2.5: OLS regression Trust Game by contribution

Dep. variable: Amount sent to receiver	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(3)	(4)	(5)	(6)
Shona(=1, Ndebele=0)	-0.192	0.000	0.403*	0.273	-0.265*	-0.532**
	(0.116)	(0.367)	(0.178)	(0.310)	(0.129)	(0.176)
Beliefs	0.623***	0.596***	0.647***	0.647***	0.692***	0.691***
	(0.035)	(0.038)	(0.026)	(0.025)	(0.055)	(0.062)
Self-identity (Base = African)						
Ndebele		0.547		0.399*		-0.199
		(0.359)		(0.194)		(0.303)
Shona		-0.761*		-0.473		0.200
		(0.401)		(0.389)		(0.303)
Zimbabwean		-0.004		0.154		-0.126
		(0.186)		(0.281)		(0.271)
Polarisation						

Table 2.5 – continued from previous page

	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(3)	(4)	(5)	(6)

Shona align		0.174		-0.002		0.016
		(0.250)		(0.132)		(0.139)
Ndebele align		0.319		-0.111		-0.085
		(0.175)		(0.111)		(0.108)
Proud		-0.136		0.389**		0.123
		(0.198)		(0.141)		(0.091)
Controls	NO	YES	NO	YES	NO	YES
Constant	2.134***	1.263	1.092**	-0.100	2.291***	3.474***
	(0.283)	(0.846)	(0.415)	(1.049)	(0.404)	(0.660)
N	344	303	344	303	344	303

Note: Standard errors clustered at the subject level in parentheses.*** p<0.01,** p<0.05,* p<0.10.

2.5.4 Public Goods Game

Figure 2.17 shows the beliefs and the actual receivers. We find that participants' beliefs of other Zimbabwean and Shona players are 5.82 and 5.46, respectively, slightly lower than the amount sent. In comparison, the amount sent to Zimbabwean and Shona receivers are 5.86 and 5.51. In contrast, beliefs about Ndebele players are slightly greater than the amount sent (5.11 and 5.026, respectively). In Figure 2.18, zero indicates the mean contributions for all receivers. Compared to mean contributions, Shona participants contribute less to Zimbabwean and Ndebele receivers. However, Shona participants contribute more than the average contributions to the in-group. Whereas Ndebele participants contribute more than the mean level of contributions to Zimbabwean (mixed group) and Ndebele (in-group) receivers, but less for Shona receivers (out-group) [11].

Again, we also examine differences in contributions by ethnic identity using Wilcoxon rank-

[11]Throughout the games, we observe high levels of contributions. Gächter & Renner (2010) find that there are cultural differences in the amount of contributions in the PGG at the aggregate level. In addition, Cardenas & Carpenter (2008) compares trust games conducted with student and non-student participants worldwide and finds a range of amounts sent from 0.30 to 0.73, mean equal to 0.51 and a standard deviation of 0.10. Our trust game findings are within the standard deviation of the mean. Kohler (2013) find high levels of transfers in experiments conducted in Zimbabwe.

sum test. In the public goods game, we find statistically different levels of cooperation across all choice contexts - as the identity of the three receivers varies - by the two ethnic groups (p=0.002, p=0.099, p=0.002). Hence, we find evidence of taste-based discriminatory behaviour between the two ethnic groups across all receiver identities.

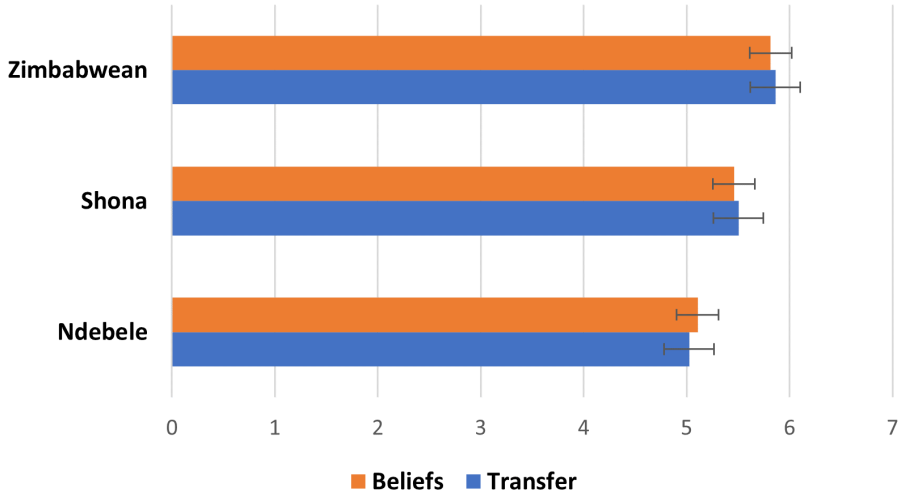


Figure 2.17: Public Goods Game (Mean contributions and beliefs)

Note: The figure shows the amount the sender chooses to send and the beliefs of the corresponding receiver. The orange bar shows beliefs, and the blue bars present the amount sent for each receiver (Zimbabwean, Shona and Ndebele). Error bars indicate 95% confidence intervals.

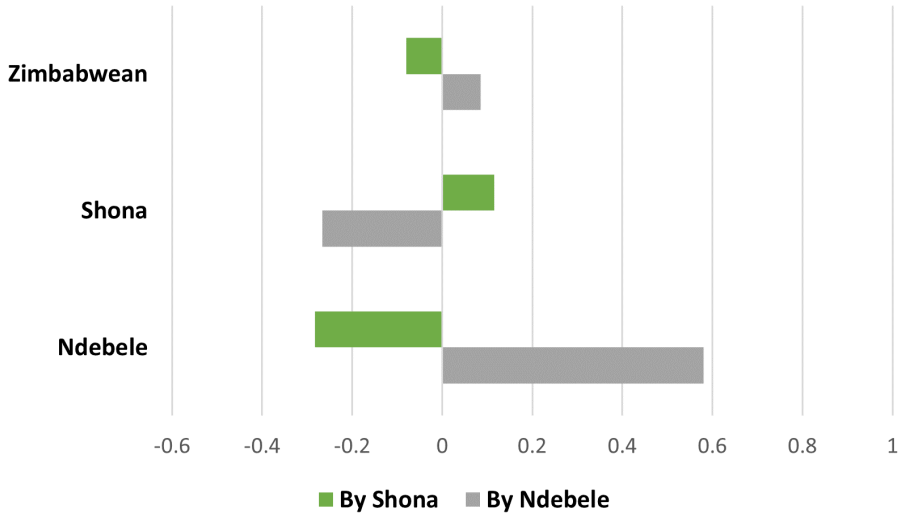


Figure 2.18: Public Goods Game (mean-centred)

Note: Zero shows the mean as the benchmark. The orange bars shows differences from the mean points sent by Shona participants. The blue bars show differences from the mean points sent by Ndebele participants.

Table 2.6 presents OLS model estimates of cooperation on identity and the intervention, including additional controls, polarisation, sex, employment status, age, household size, marital status and risk. In columns (1) and (2), the dependent variable is the amount of cooperation with Ndebele receivers. In columns (3) and (4), the dependent variable is the amount of cooperation with Shona receivers. Lastly, in columns (5) and (6), the dependent variable is the amount of cooperation with Zimbabwean receivers. Shona participants send more than Ndebele dictators when the point goes to Shona and Zimbabwean receivers. However, this result is not robust to the inclusion of controls. Beliefs significantly influence behaviours. The beliefs regarding how Ndebele and Shona people would behave significantly increase contributions for all receivers.

Self-identifying as Ndebele reduces points sent to Shona receivers. In addition, self-identifying as Zimbabwean reduces points sent to Shona and Ndebele receivers. Qualitatively we observe positive transfers to Zimbabwean receivers, although this is not significantly different. The self-reported polarisation measures reveal that being highly aligned to the Shona identity decreases points sent to Ndebele and Zimbabwean receivers. Being proud of one's identity increases points sent to Shona receivers.

Table 2.6: OLS regression Public Goods Game by receiver

Dep. variable: Amount sent to receiver	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(3)	(4)	(5)	(6)
Shona(=1, Ndebele=0)	-0.342	-0.214	0.332**	0.057	0.395*	0.136
	(0.238)	(0.285)	(0.144)	(0.550)	(0.211)	(0.329)
Beliefs	0.631***	0.626***	0.645***	0.645***	0.699***	0.685***
	(0.052)	(0.046)	(0.032)	(0.034)	(0.018)	(0.017)
Self-identity (Base = African)						
Ndebele		-0.288		-0.908**		-0.855
		(0.509)		(0.388)		(0.479)
Shona		-0.410		0.505		0.735
		(0.331)		(0.601)		(0.822)
Zimbabwean		-0.426***		-0.371**		0.136

Table 2.6 – continued from previous page

	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(3)	(4)	(5)	(6)

		(0.123)		(0.157)		(0.086)
Polarisation						
Shona align		-0.264*		-0.280		-0.199
		(0.141)		(0.194)		(0.149)
Ndebele align		-0.082		-0.184		-0.194
		(0.111)		(0.149)		(0.156)
Proud		-0.078		0.292*		-0.047
		(0.124)		(0.142)		(0.129)
Controls	NO	YES	NO	YES	NO	YES
Constant	2.609***	2.772***	1.438***	2.420	1.098***	2.265**
	(0.463)	(0.729)	(0.214)	(1.397)	(0.292)	(0.937)
N	314	303	314	303	314	303

Note: Standard errors clustered at the subject level in parentheses.*** p<0.01,** p<0.05,* p<0.10.

In sum, we find evidence of in-group favouritism and some evidence of out-group discrimination. When given the option between several identities, participants who chose to identify as Ndebele, Shona, and Zimbabwean were less cooperative with out-group members, respectively. Regarding the polarisation measures, we observe that the multi-player game - the public goods game - creates greater importance of the self-reported polarisation measures. [12].

2.5.5 Social norms elicitation task

We first report the evaluations of actions and test for differences in evaluations between Ndebele and Shona participants based on the reference group. From this, we further examine the impact of social norms on discriminatory behaviour is investigated. Lastly, differences between individual and group norms are tested to identify the mechanism of polarisation where the norms of the in-group become more extreme and opposed to private attitudes.

[12]We also estimated the same set of regressions using session fixed effects rather than individual effects. The results are unchanged. We also estimated the same set of regressions using Tobit estimates to account for censoring. The results are relatively unchanged, see Appendix A.5.1, Table A1 - Table A3

We take a similar approach to Krupka & Weber (2013). To begin, we converted participants' responses into numerical scores. Participants who evaluate an action to be "very socially inappropriate" is valued at -1, "somewhat socially inappropriate" is valued at -1/3, "somewhat socially appropriate" is valued at 1/3 and "very socially appropriate" is valued at 1. Table 2.7 - Table 4.6 presents participants evaluations of social appropriateness. We represent the findings by the majority action. The first three rows present the total evaluation of social norms by all participants. The following rows divide the sample by the identity of the evaluator.

In the Pay and Tax norms scenarios, the modal evaluations are similar overall and by the two ethnicities, see Table 2.7 and Table 2.8. The perceptions of nepotism and tax evasion are evaluated as *very socially inappropriate*. We use Wilcoxon rank-sum to compare the evaluations between the two ethnicities. Testing the pay scenario reveals that evaluations between Shona and Ndebele participants differ for evaluations of the Ndebele majority, Shona majority and own perceptions ($p = 0.001$, $p = 0.004$, and $p = 0.002$, respectively). Comparing differences in evaluations between the Ndebele and Shona following the tax scenario, we find that perceptions differ between the two groups regarding the modal response of Ndebele people ($p = 0.003$) and their perceptions ($p = 0.039$). However, we find no difference in evaluations between the two groups regarding the majority of Shona people's willingness to falsely report sales to pay less tax. Although modal responses are similar in these scenarios, non-parametric tests reveal that norm evaluations of the majority of Ndebele and Shona people differ between participants.

In the Family scenario, we observe differences in perceptions when the player is Shona. See Table 3.10. Shona players view that the majority of Ndebele would perceive marriage between the two ethnic groups as being *somewhat socially inappropriate*. In contrast, their in-group would view this as *somewhat socially appropriate*. Lastly, their perceptions regarding this scenario are evaluated as *very socially appropriate*. On the other hand, Ndebele participants perceive that the majority of Shona people and Ndebele people view this scenario as *very socially appropriate*. Similarly, their perceptions are that this scenario is *very socially appropriate*. To directly compare the differences in perceptions between the two groups for each evaluation, we use Wilcoxon rank-sum test. We find that evaluations between Shona and Ndebele participants differ for evaluations of the Ndebele majority, Shona majority and own perceptions ($p = 0.061$, $p = 0.011$, and $p = 0.061$, respectively).

Table 4.5 presents the evaluations of reporting bribery. We find that both Shona and Ndebele players evaluate that most Ndebele people would view this action as *very socially appropri-*

ate. Whereas an individual's own perceptions and the majority of Shona participants view reporting bribery as *very socially inappropriate*. We use the Wilcoxon rank-sum test to test the differences in evaluations by the two ethnolinguistic groups. We find that there is no statistical difference in evaluations by the two groups across all three variations ($p = 0.680$, $p = 0.824$, and $p = 0.382$, respectively). Therefore, evaluations are consistent across identities.

In the scenario where the action is to punish bribery, we find that all evaluations view punishing bribery as *very socially appropriate*. To directly compare the differences in perceptions between the two groups for each evaluation, we use Wilcoxon rank-sum test. We find that there is no statistical difference in evaluations by the two groups across all three variations ($p = 0.116$, $p = 0.360$, and $p = 0.554$, respectively). Therefore, we observe consistent evaluations of others by all players in the two bribe scenarios.

Table 2.7: Elicited social appropriateness of ethnic favouritism leading to higher pay (Pay scenario)

	−−	−	+	++
Ndebele majority	54.59	20.24	13.65	11.53
Shona majority	48.47	20.71	18.35	12.47
Own perception	67.53	16.00	7.53	8.94
Participant is Ndebele				
Ndebele majority	66.01	16.99	9.15	7.84
Shona majority	38.56	22.22	27.45	11.76
Own perception	77.12	12.42	5.23	5.23
Participant is Shona				
Ndebele majority	48.88	22.01	15.30	13.81
Shona majority	54.48	19.78	12.69	13.06
Own perception	62.69	17.54	8.58	11.19

Note: Responses are "very socially inappropriate" (–), "somewhat socially inappropriate" (-), "somewhat socially appropriate" (+), "very socially appropriate" (++). Modal responses are shaded.

To identify divergence between the in-group and their own group as an indicator of more extreme group norms, as opposed to personal norms, in other words, pluralistic ignorance (Smerdon et al. 2020, Mackie 1986). We test differences in group norms and individual norms. In the pay scenario, we find that for Ndebele participants, across all treatments, there is no evidence of divergence between personal and group norms. However, we observe divergence in norm evaluations when the participant is Shona and in the *Divide* treatment

Table 2.8: Elicited social appropriateness of falsely reporting sales (Tax scenario)

	−−	−	+	++
Ndebele majority	44.34	29.48	19.34	6.84
Shona majority	39.62	30.42	17.45	12.50
Own perception	63.68	21.70	10.14	4.48
Participant is Ndebele				
Ndebele majority	51.63	31.37	13.73	3.27
Shona majority	40.52	29.41	20.26	9.80
Own perception	68.63	22.88	6.54	1.96
Participant is Shona				
Ndebele majority	40.67	27.99	22.76	8.58
Shona majority	38.81	31.34	16.04	13.81
Own perception	60.82	20.90	12.31	5.97

Note: Responses are "very socially inappropriate" (–), "somewhat socially inappropriate" (-), "somewhat socially appropriate" (+), "very socially appropriate" (++). Modal responses are shaded.

Table 2.9: Elicited social appropriateness of someone marrying an individual of the opposite ethnicity (Family scenario)

	−−	−	+	++
Ndebele majority	21.09	28.67	27.49	22.75
Shona majority	12.80	22.75	33.65	30.81
Own perception	5.45	7.11	25.12	62.32
Participant is Ndebele				
Ndebele majority	22.22	20.92	26.80	30.07
Shona majority	22.22	24.18	22.22	31.37
Own perception	3.92	5.23	20.92	69.93
Participant is Shona				
Ndebele majority	20.52	32.84	27.99	18.66
Shona majority	7.46	22.01	39.93	30.60
Own perception	6.34	8.21	27.61	57.84

Note: Responses are "very socially inappropriate" (–), "somewhat socially inappropriate" (-), "somewhat socially appropriate" (+), "very socially appropriate" (++). Modal responses are shaded.

($p = 0.021$ t-test). In the tax scenario, both Ndebele and Shona participants differ in their perceptions in the control treatment only. We observe divergence in norms in all treatments in the family scenario for both Ndebele and Shona participants ($p < 0.05$). No difference is

Table 2.10: Elicited social appropriateness of reporting bribery (Bribery1 scenario)

	−−	−	+	++
Ndebele majority	24.52	25.00	18.10	32.38
Shona majority	30.57	22.51	24.64	22.27
Own perception	40.62	10.93	12.11	36.34
Participant is Ndebele				
Ndebele majority	26.14	21.57	16.99	35.29
Shona majority	29.41	24.84	25.49	20.26
Own perception	43.79	9.15	12.07	33.99
Participant is Shona				
Ndebele majority	23.68	27.07	18.42	30.83
Shona majority	31.34	21.27	23.88	23.51
Own perception	38.95	11.99	11.24	37.83

Note: Responses are "very socially inappropriate" (–), "somewhat socially inappropriate" (-), "somewhat socially appropriate" (+), "very socially appropriate" (++). Modal responses are shaded.

Table 2.11: Elicited social appropriateness of punishing bribery (Bribery2 scenario)

	−−	−	+	++
Ndebele majority	8.65	12.74	29.09	49.52
Shona majority	7.66	14.35	28.47	49.52
Own perception	8.85	5.50	23.68	61.96
Participant is Ndebele				
Ndebele majority	7.84	12.42	24.18	55.56
Shona majority	7.84	11.76	28.10	52.29
Own perception	9.15	8.50	16.99	65.36
Participant is Shona				
Ndebele majority	9.13	12.93	31.94	46.02
Shona majority	7.55	15.85	23.68	47.92
Own perception	8.68	3.77	27.55	60.00

Note: Responses are "very socially inappropriate" (–), "somewhat socially inappropriate" (-), "somewhat socially appropriate" (+), "very socially appropriate" (++). Modal responses are shaded.

found in the first bribery scenario across identities and treatments. In the second bribery scenario, we find both Ndebele and Shona participants differ in perceptions in the *control* treatment ($p = 0.061$ and $p = 0.000$, respectively).

In Table 2.12 the dependent variable is mean contributions pooled over all receivers for the prisoner's dilemma (see column (1)), dictator game (see column (2)), trust game (see column (3)) and public goods game (see column (4)). The more socially appropriate marriage between ethnicities is evaluated, the higher likelihood of coordination and cooperation in the prisoner's dilemma and dictator game. If participants are Shona, the evaluations of the social appropriateness of interethnic marriage reduce coordination and cooperation in the prisoner's dilemma and dictator game. We find that in the public goods game when participants view the pay scenario (nepotism) as socially appropriate leads to less mean cooperation. However, if the participant is Shona, we find evidence of the inverse relationship. This result suggests that when the participant is Shona, the more socially appropriate nepotism is evaluated, the greater the contribution to the common pot in the public goods game. Evaluating reporting bribery as socially appropriate leads to greater cooperation in columns (2) to (4). However, if the participant evaluating the norm is Shona, reporting bribery reduces cooperation in the dictator game. Punishing bribery reduces coordination and cooperation in the prisoner's dilemma and the dictator game. Interpreting the marginal effects of the interaction term, punishing bribery when the evaluator is Shona also reduces cooperativeness in the prisoner's dilemma and dictator game, although the negative effect on cooperation is larger for Shona participants. Hence, we find that social norms shape overall cooperative behaviour.

Table 2.12: OLS Regression on effects of social norms on mean contributions in games

	PD	DG	TG	PGG
	(1)	(2)	(3)	(4)
Shona	-0.157	-0.858	-0.527	0.788
	(0.119)	(0.458)	(0.648)	(0.826)
Family	0.183***	0.769***	0.267	0.074
	(0.018)	(0.199)	(0.230)	(0.101)
Family x Shona	-0.200***	-1.118**	-0.146	0.011
	(0.032)	(0.398)	(0.286)	(0.108)
Pay	0.042	-0.028	-0.548	-0.960***
	(0.077)	(0.581)	(0.324)	(0.179)
Pay x Shona	0.009	-0.021	0.262	0.868*
	(0.077)	(0.756)	(0.508)	(0.430)
Tax	-0.072	0.274	-0.076	-0.288

Continued on next page

Table 2.12 – continued from previous page

	PD	DG	TG	PGG
	(1)	(2)	(3)	(4)
	(0.055)	(0.748)	(0.576)	(0.152)
Tax x Shona	0.081	-0.353	-0.229	0.494
	(0.083)	(0.961)	(0.620)	(0.298)
Bribery1	0.037	1.130***	0.416*	0.350***
	(0.051)	(0.224)	(0.196)	(0.032)
Bribery1 x Shona	-0.044	-1.214**	-0.281	0.068
	(0.051)	(0.470)	(0.281)	(0.193)
Bribery2	-0.116**	-1.020***	-0.787	-0.162
	(0.038)	(0.213)	(0.442)	(0.092)
Bribery2 x Shona	0.194**	0.818**	0.322	-0.107
	(0.072)	(0.243)	(0.436)	(0.410)
Controls	YES	YES	YES	YES
Constant	0.779***	3.269*	2.089	1.449
	(0.182)	(1.610)	(1.441)	(1.570)
N	264	190	190	190

Note: Standard errors clustered at the subject level in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

To sum up, the evaluations of the social norms suggest that there are diverging norms between the majority of Ndebele people and the majority of Shona people regarding actions surrounding reporting bribery. Similarly, when evaluating the social appropriateness of interethnic marriages. However, both nepotism in the workplace leading to greater pay and falsely reporting sales to pay less tax are evaluated as *very socially inappropriate* across all reference groups. Our regression analysis results show differences in normative evaluations and subsequent cooperative behaviours in different scenarios. We find evidence of pluralistic ignorance through differences between personal and group norms.

2.5.6 Social norm intervention

We now turn to the findings from the social norm intervention, which occurred in the second part of the experiments, to identify the extent that behaviours can be subtly shifted to yield greater common good (greater cooperation, trust and altruism) or foster less prosocial behaviours (further intergroup discrimination). In separate treatments, participants receive an injunctive social norm, either *Heal* or *Divide* or no information. No information is provided in the *Control* treatment, which acts as a baseline for measuring the susceptibility to social influence when informed of the prevailing social norms. In either treatment, participants then play the games and the social norms elicitation task, from which the social norms intervention is derived. We begin by identifying the impact of the intervention on behaviours. Afterwards, the analysis considers the impacts of the intervention on the social norms following the information provided.

Figure 2.19a displays the main findings of the prisoner's dilemma following the interventions [13]. In the *Control* group, 47.1% of participants chose the Pareto optimal outcome of coordination on average for all receivers. In the group that received information on shared norms - *Heal* treatment - 63.6% of participants coordinated across all receivers. Following the *Divide* treatment, 70.4% of participants chose the Pareto optimum outcome of coordination when pooled across all receivers. We conduct Wilcoxon rank-sum tests to test whether coordination differs statistically between treatments. We find that coordination in both the *Heal* and *Divide* treatments are statistically different from the *Control* (p=0.000 and p=0.000). Therefore, in Figure 2.19a, we observe that the interventions both increase prosocial behaviour compared to receiving no information.

Probit regression analysis of the intervention on coordination, where the base category is the *Control* group of the social norms intervention and Ndebele participants, see Table 2.13, Panel A. The results show that the *Heal* reduces coordination with Ndebele receivers but increases coordination with Shona participants. The *Divide* treatment increases coordination with Zimbabwean receivers. In column (3), we include the interaction term to assess the effects of the intervention depending on the participant's identity. The *Divide* treatment increases coordination with Shona receivers when the dummy variable *Shona* equals zero for Ndebele participants. Conversely, if the participant is Ndebele, the *Divide* treatment increases coordination with Zimbabwean receivers [14]. Hence, consistent with our hypothesis,

[13] See Appendix A.5.2, Figure A.7 - Figure A.10 for disaggregated bar charts for each receiver by treatment.

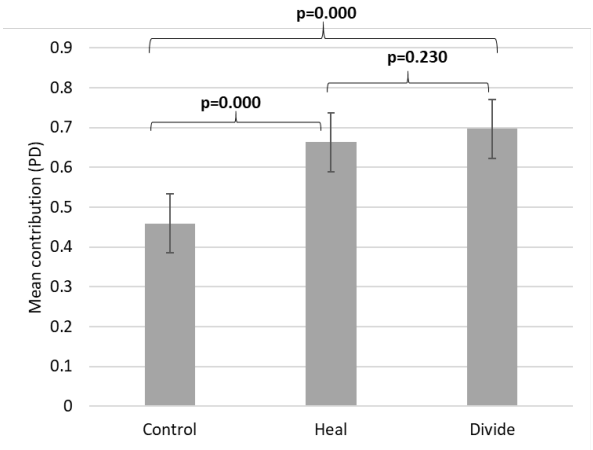
[14] We cannot directly report the results of interaction terms between the treatments and identity indicator due to collinearity.

the *Heal* treatment bridges the gap between groups and reduces intergroup discrimination. However, the *Divide* treatment has the opposite effect and increases coordination towards non-rival receivers.

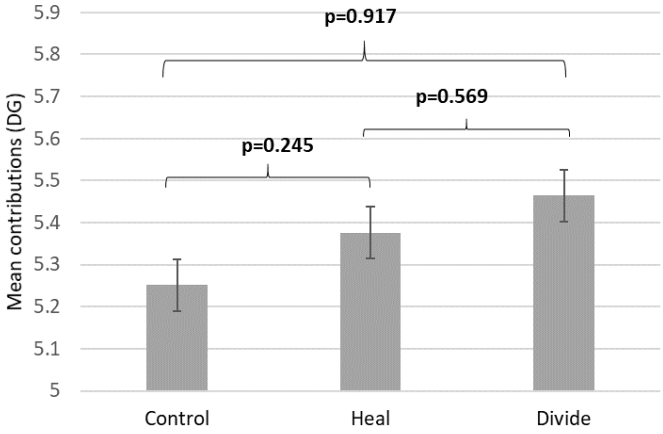
Figure 2.19b presents mean contributions in the dictator game by intervention. In the *control* group, dictators transfer 5.25 points on average for all receivers. Following the *Heal* treatment, dictators transfer, on average, 5.38 points. Participants that received information of diverging social norms - *Divide* treatment - transfer 5.46 points. We observe that transfers are slightly greater in the *Heal* and still greater in the *Divide* treatment compared to the *Control*. However, there is no statistically significant difference following the treatments compared to the baseline.

Table 2.13, Panel B presents estimated OLS regression of contributions, including the intervention, identity, beliefs and individual-level characteristics in the dictator game. The *Heal* treatment increases transfers compared to the *Control* when the contributions go to a Ndebele receiver and a Shona receiver, when the dummy variable *Shona* equals zero, therefore, for Ndebele participants. Moreover, the *Divide* treatment increases transfer for Shona and Zimbabwean receivers. Interpreting the marginal effects of the interaction term, if the dictator is Shona, the *Heal* increases cooperation with Ndebele receivers. However, the effect is smaller than the change when the participant is Ndebele. However, following the *Divide* treatment, Shona dictators transfer significantly less to Zimbabwean receivers. Hence, the *Heal* treatment increases cooperation between Ndebele players and Shona receivers. Conversely, the *Divide* treatment has the greatest effect of reducing cooperation with the non-rival Zimbabwean receiver.

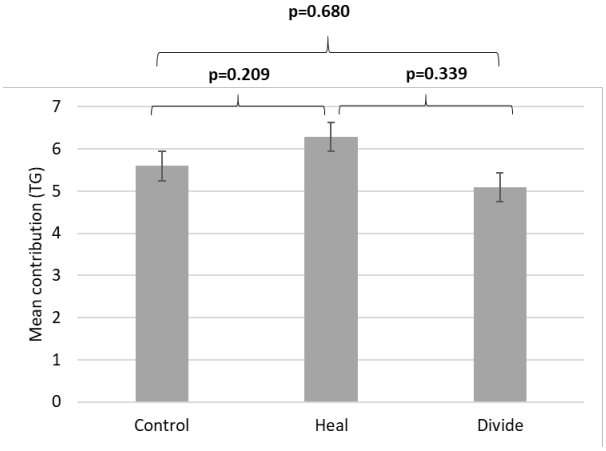
Figure 2.19c presents the mean contributions in the trust game by intervention. In the baseline treatment, participants transfer 5.59 points on average. Participants in the *Heal* treatment transfer 6.28 points. Lastly, following the *Divide* treatment, participants transfer 5.09 points. However, we do not find a statistically significant difference following either the *Heal* or *Divide* in comparison to the *Control*, respectively.



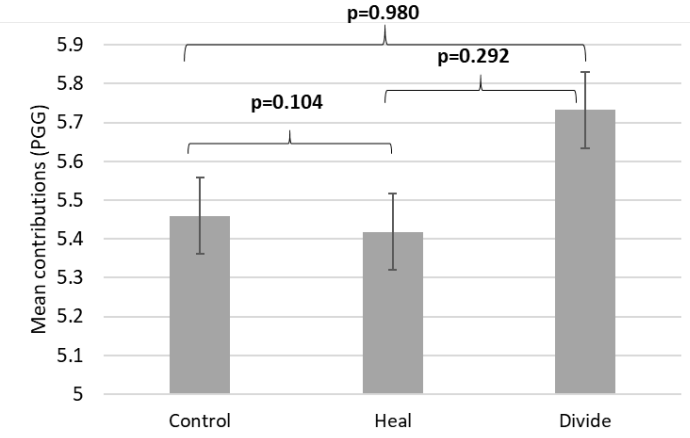
(a) The figure shows the fraction of coordination by the intervention. Error bars indicate 95% confidence intervals.



(b) The figure shows the mean contribution in the Dictator Game by the intervention treatment. Error bars indicate 95% confidence intervals.



(c) The figure shows the mean contribution in the Trust Game by the intervention treatment. Error bars indicate 95% confidence intervals.



(d) The figure shows the mean contribution in the Public Goods Game by the intervention treatment. Error bars indicate 95% confidence intervals.

Table 2.13: Intervention effects on social norms

Dep. variable: Amount sent to receiver	Ndebele			Shona			Zimbabwean		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Prisoners Dilemma									
Heal	-0.577***	-0.503**	-0.174	1.397***	1.687***	1.540***	0.075	0.109	0.199
	(0.151)	(0.231)	(0.343)	(0.180)	(0.315)	(0.370)	(0.111)	(0.218)	(0.255)
Divide	0.175	0.095	-0.082	0.285**	0.216	-0.236	0.515***	0.759***	0.734***
	(0.142)	(0.237)	(0.248)	(0.118)	(0.192)	(0.298)	(0.192)	(0.193)	(0.195)
Shona			-0.005			0.085			-0.083
			(0.148)			(0.218)			(0.201)
Heal x Shona			-0.407			0.000			-0.105
			(0.394)			(.)			(0.393)
Divide x Shona			0.000			0.550			0.000
			(.)			(0.345)			(.)
Controls	NO	YES	YES	NO	YES	YES	NO	YES	YES
Constant	-0.222	-1.753	-1.700	0.480***	-1.126	-1.005	0.529***	-0.436	-0.516
	(0.146)	(1.324)	(1.257)	(0.122)	(0.931)	(1.055)	(0.095)	(0.365)	(0.529)
N	288	266	263	288	266	262	288	254	253
Panel B: Dictator Game									
Heal	0.216	0.510	1.461***	0.182	0.683***	1.051**	-0.420	0.137	-0.373
	(0.588)	(0.419)	(0.201)	(0.250)	(0.160)	(0.452)	(0.389)	(0.144)	(0.556)
Divide	0.385	0.393	1.229	0.451*	0.642	1.378**	-0.013	0.096	1.328**

Continued on next page

Table 2.13 – continued from previous page

Dep. variable: Amount sent to receiver	Ndebele			Shona			Zimbabwean		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
	(0.541)	(0.308)	(0.667)	(0.209)	(0.358)	(0.491)	(0.353)	(0.491)	(0.414)
Shona			-0.524			0.591*			0.515**
			(0.310)			(0.285)			(0.174)
Heal x Shona			-1.104*			-0.436			0.561
			(0.516)			(0.496)			(0.600)
Divide x Shona			-0.955			-0.828			-1.360*
			(0.734)			(0.734)			(0.594)
Controls	NO	YES	YES	NO	YES	YES	NO	YES	YES
Constant	4.520***	2.550**	1.970*	5.026***	0.230	0.817	5.703***	0.610	1.206
	(0.413)	(1.066)	(0.936)	(0.208)	(0.854)	(0.811)	(0.296)	(1.215)	(1.253)
N	368	303	303	368	303	303	368	303	303
Panel C: Trust Game									
Heal	0.194	0.114	0.653*	0.350	0.028	1.204***	-0.029	-0.289	-0.157
	(0.532)	(0.186)	(0.336)	(0.234)	(0.209)	(0.214)	(0.485)	(0.400)	(0.230)
Divide	0.858	0.702*	1.835***	0.730*	0.761***	3.041***	0.290	0.031	1.464***
	(0.695)	(0.329)	(0.538)	(0.369)	(0.207)	(0.270)	(0.383)	(0.132)	(0.122)
Heal x Shona			-0.639			-1.388***			-0.179
			(0.344)			(0.274)			(0.522)
Divide x Shona			-1.274**			-2.567***			-1.598***
			(0.437)			(0.386)			(0.125)

Continued on next page

Table 2.13 – continued from previous page

Dep. variable: Amount sent to receiver	Ndebele			Shona			Zimbabwean		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Controls	NO	YES	YES	NO	YES	YES	NO	YES	YES
Constant	5.071***	1.663*	1.720**	5.461***	0.348	0.751	6.067***	3.554***	3.118***
	(0.444)	(0.823)	(0.639)	(0.218)	(1.006)	(0.928)	(0.331)	(0.659)	(0.627)
N	364	303	303	364	303	303	364	303	303
Panel D: Public Goods Game									
Heal	0.037	-0.083	-0.089	0.208	0.082	-0.100	-0.054	-0.317	-0.721***
	(0.522)	(0.277)	(0.262)	(0.334)	(0.279)	(0.432)	(0.336)	(0.430)	(0.208)
Divide	0.722	0.214	0.268	0.433	0.366	-0.074	0.087	-0.227	0.461
	(0.667)	(0.353)	(0.579)	(0.376)	(0.197)	(0.437)	(0.274)	(0.469)	(0.305)
Shona			-0.219			0.027			0.129
			(0.280)			(0.516)			(0.337)
Heal x Shona			0.006			0.217			0.449
			(0.534)			(0.719)			(0.630)
Divide x Shona			-0.060			0.494			-0.756
			(0.711)			(0.543)			(0.728)
Controls	NO	YES	YES	NO	YES	YES	NO	YES	YES
Constant	4.944***	2.920***	2.703***	5.465***	2.624	2.650*	5.784***	2.196**	2.380**
	(0.388)	(0.701)	(0.493)	(0.198)	(1.429)	(1.154)	(0.244)	(0.860)	(0.769)
N	360	303	303	360	303	303	360	303	303

Continued on next page

Table 2.13 – continued from previous page

Dep. variable: Amount sent to receiver	Ndebele			Shona			Zimbabwean		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)

Note: Regression analysis of amounts sent to each receiver. In column (2), we add demographic controls, and in column (3), we add the interaction term on the treatments. Robust standard errors clustered by the session in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

To further test the impacts of the intervention on behaviour, Table 2.13, Panel C, displays the OLS regression model of behaviour in the trust game on identity, intervention and a set of controls. Interpreting the interaction terms in columns (3), Ndebele participants in the *Heal* treatment are more trusting of the in-group and out-group. However, Shona trustors in the *Heal* treatment contribute less to their in-group, therefore, causing in-group discrimination and thus the norm-nudge has had an unintended consequence[15]. Ndebele participants in the *Divide* treatment are more trusting towards all receivers. Shona trustors in the *Divide* treatment are more trusting toward out-group and in-group receivers. However, they are less trusting toward Zimbabwean receivers.

Figure 2.19d presents the mean contributions in the public goods game by intervention. On average, participants in the control group transfer 5.46 points to the communal pot. Those in the *Heal* treatment transfer 5.42 points, and those in the *Divide* transfer 5.73 points. The transferred points vary between treatments compared to the *Control*. However, we do not find a statistically significant difference in mean transfers over all receivers. Table 2.13, Panel D. we further investigate the impacts of the interventions on transfers in the public goods game. The *Heal* treatment reduces cooperation with Zimbabwean players when the participant is Ndebele. We find no other evidence of treatment effects influencing behaviour in the public goods game [16].

Table 2.14 reports the results of regression models that investigate if the social norms intervention influences reported social norms (Miller & Prentice 2016). In column (1), we run an OLS regression model, where the base category is the *control* group of the social norms intervention. Given the ordinal nature of the modal evaluations, we use ordered Probit regression model in column (2). The dependent variable measures the appropriateness of actions. We find that the *Heal* intervention, when the dummy variable *Shona* equals zero, therefore, Ndebele participants, increase evaluations of the social appropriateness in the Family Scenario. Conversely, when the participant is Shona, the *Heal* treatment lowers social appropriateness evaluations. For Ndebele participants, the *Divide* treatment reduces ratings

[15](see, (Dolan et al. 2012) and (Bolton et al. 2019) for review of cases when nudges backfire)

[16]Disaggregated analysis, separated by identity yields qualitatively similar results, see Appendix A.5.2, Table D1 - Table D4

in this scenario. Therefore, the intervention directly alters the scenario used. This shows evidence of consistent majoritarian behaviour, given the similarities between the Family scenario and the treatments. However, for Shona evaluators, the *Divide* treatment increases the appropriateness of interethnic marriages.

Overall, we find mixed results of the influence of the social norm-based intervention. We find that the intervention does not largely impact the pooled behaviour across all receivers. However, we do find that the identity of the receiver matters. The regression analysis reveals that Ndebele participants generally respond positively to the *Heal* and *Divide* treatments, increasing cooperation for in-group, out-group and Zimbabwean receivers. However, we find that the *Heal* treatment backfires in the public goods game by reducing cooperation from Ndebele participants to Zimbabwean receivers.

On the other hand, when the participant is Shona, the *Heal* treatment increases out-group cooperation in the dictator game - in line with our hypothesis. However, the *Heal* treatment backfires and causes in-group discrimination between Shona receivers and senders. In addition, the *Divide* treatment reduces the amount sent from Shona players when the receiver is Zimbabwean. Therefore, the *Divide* treatment negatively impacts behaviour with the non-rival receiver.

Table 2.14: Intervention effects on social norms

Dep. variable: Norm evaluations	Family		Pay		Tax		Bribery 1		Bribery 2	
Model	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Shona	-0.151	-0.322*	0.327***	0.638***	0.283**	0.511**	-0.161	-0.268	-0.226*	-0.525**
	(0.083)	(0.169)	(0.066)	(0.130)	(0.121)	(0.221)	(0.109)	(0.208)	(0.114)	(0.214)
Heal	0.296*	0.762**	0.129	0.206	0.176	0.340	0.028	-0.067	0.157	0.342
	(0.130)	(0.316)	(0.085)	(0.172)	(0.226)	(0.458)	(0.114)	(0.190)	(0.091)	(0.317)
Heal x Shona	-0.212	-0.591*	-0.085	-0.140	-0.185	-0.309	0.053	0.193	-0.164*	-0.391
	(0.138)	(0.342)	(0.100)	(0.214)	(0.236)	(0.476)	(0.107)	(0.186)	(0.075)	(0.284)
Divide	-0.680***	-1.421***	0.337***	0.554***	-0.002	-0.083	0.335***	0.500***	0.052	0.213
	(0.071)	(0.194)	(0.047)	(0.130)	(0.099)	(0.198)	(0.080)	(0.138)	(0.051)	(0.149)
Divide x Shona	0.552***	1.129***	-0.399***	-0.703***	-0.097	-0.125	-0.193	-0.254	0.044	0.103
	(0.104)	(0.263)	(0.082)	(0.155)	(0.164)	(0.308)	(0.106)	(0.165)	(0.090)	(0.321)
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	0.351		-0.682***		0.604		-0.192		0.750***	
	(0.229)		(0.191)		(0.366)		(0.276)		(0.218)	
N	303	303	303	303	303	303	303	303	301	301

Note: Column (1), we run an OLS regression and in column (2), we run an Ordered Probit regression model. Robust standard errors clustered by session in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

2.6 Discussion and conclusion

This study aimed to investigate in-group favouritism and out-group discrimination using ethnolinguistic groups in Zimbabwe. The long-run political conflict entangled in ethnic ties between the Ndebele and Shona people provides a pertinent example of group affiliation that can foster group polarisation and, subsequently, discriminatory behaviour. Polarisation can arise out of focus on the external other, but it reflects or draws upon internal us vs them dynamic (McCoy et al. 2018). That can impact social cohesion and the government's ability to execute policies (McCoy et al. 2018). We set out to identify discriminatory behaviour using four experimental games, which include an in-group, out-group and neutral (Zimbabwean) receiver. Participants played as the sender in various experimental games; we also elicited beliefs. Therefore, participants also play the role of the receiver. Depending on the receiver's identity, taste-based discrimination can be measured by differences in transfers by participants. In addition, the influence and importance of beliefs identify statistical discrimination. In line with social identity theory (Tajfel et al. 1971, Akerlof & Kranton 2000, 2005), we proposed that participants will favour their in-group and want their in-group to be better off than the out-group.

In sum, we find evidence of discriminatory behaviour between the two ethnic groups, as Shona participants behave more prosocially towards the in-group and less towards the out-group. Alongside identifying taste-based and statistical discrimination, this study aimed to assess the perceptions around the identities that can constitute polarised views. Contrary to our conceptualisation of polarisation which takes from Esteban & Ray (1994) and Abbink & Harris (2019)'s notion of non-rival groups. In general, there are significant differences in behaviours towards the 'Zimbabwean' or non-rival group receiver. The research also identifies statistical discrimination, as participants were pessimistic about the amount Ndebele participants would send in return across all game forms. We find that beliefs about others significantly influence prosocial behaviours, often more important than actual identity. Perhaps the narrative used by the government that emphasised differences between Ndebele and Shona people has influenced beliefs, with incomplete information supporting statistical discrimination or stereotyping (Bonomi et al. 2021).

We also elicit social norms using real-world scenarios. From the social norms elicitation task, we observe in both the family scenario and Bribe Scenario 1 the idea of pluralistic ignorance due to the difference in their perception compared to the modal perception (Smerdon et al. 2020, Miller & Prentice 2016, Mackie 1986). It appears that Shona participants privately view interethnic marriages as socially inappropriate but perceive most people in the in-group as viewing this as socially appropriate. While Ndebele participants privately deviate from the norm in reporting bribery scenarios. Pluralistic ignorance has been associated with negative behaviours (Smerdon et al. 2020, Miller & Prentice 2016, Bursztyn et al. 2020). For example, Ndebele people may abstain from reporting bribery to conform to the norm and potentially allow corrupt behaviours to occur.

Behaviourally informed policies can be used in the face of these societal divisions, as individual-level behaviours play a role at the community-level, social level and political sphere (Ewert & Loer 2021, Ostrom 1998). Therefore, we also set out to investigate to what extent social norm-based nudges can cultivate greater inter-group cooperation and trust. Our experiments were set out as follows: participants were randomly assigned into a *Control* group, *Heal* or *Divide* treatment. We hypothesised that following information on the shared norm (*Heal*), we will observe greater inter-group cooperation. Conversely, participants exposed to the treatment where they are informed of diverging norms (*Divide*) will show greater levels of out-group discrimination. Following the norm-based intervention, we also find that it is easier to *Divide* than to *Heal* social norm evaluations. Our finding is consistent with the literature that acknowledges that proscriptive norms garner greater behavioural responses than prescriptive (Farrow et al. 2017, Cialdini et al. 2006).

Moreover, the interaction terms reveal that Shona participants in the *Divide* treatment transfer less to the Zimbabwean receiver. Conversely, when the participant is Ndebele, the *Divide* treatment increases the amount sent to their in-group, the out-group and Zimbabwean groups. On the other hand, the *Heal* treatment bridges that gap between Ndebele and Shona participants. However, the treatment leads to lowered trust between Shona in-group members and reduces cooperation with Zimbabwean receivers by Ndebele participants. In addition, the interventions do impact norm evaluations (Dimant et al. 2020). Therefore, the treatments inconsistently alter behaviours. We observe our social norms intervention "backfire", from the intended outcomes (Acemoglu & Jackson 2017, Bolton et al. 2019, Hummel & Maedche 2019, Bicchieri & Dimant 2019, Holzmeister et al. 2022, Sunstein 2017) at times are proscriptive (negative) norm yield greater cooperation. Therefore, our experiments reveal that information regarding norms changes behaviour, although inconsistently. However,

the mixed results of the intervention may be due to issues surrounding norm-based nudges. The successfulness of norm-based nudges can depend on the beliefs and credibility of the source (Mol et al. 2020). The literature has observed that norm-based interventions can backfire (Dolan et al. 2012). The effectiveness of such norm-based policies may fail due to a general lack of confidence and trust in the source or policymaker (Tyers 2019). Therefore, policymakers have the challenge of ensuring that policies are trusted and consistent with the people's demands (Bicchieri et al. 2022, Sunstein 2021, Arad & Rubinstein 2018).

This study contributes to the growing literature on discriminatory behaviour, social norms and norm-based interventions. By altering the framing of norms, we test if we can further *Divide* or *Heal* polarised groups in Zimbabwe and find that norms shift discriminatory behaviour, although not always with the desired intent and subsequent evaluations of norms. We acknowledge that nudges should not replace traditional policies but can work alongside policies (Benartzi et al. 2017). The major limitation of this study is the number of participants allocated to each treatment group. Due to the Covid-19 pandemic, getting more participants in person was impossible. In addition, future research could consider the use of social media and online discourse as it grows in developing countries such as Zimbabwe, as Zimbabwean ethnicity and national identity continue to be discussed on social media (Mpofu 2013).

In conclusion, this paper explores discriminatory behaviour in Zimbabwe using the prisoner's dilemma, dictator game, trust game and public goods game. Based on the social identity theory, we hypothesised that there would be in-group favouritism and out-group discrimination along ethnic lines. Moreover, a shared identity group will be non-rival (Abbink & Harris 2019, Cason & Mui 1997). We also proposed that social norms and beliefs partly explain discriminatory behaviour. Our findings suggest that norm-based nudges can be used to change behaviours. However, it is not always clear in what direction this will change behaviour.

3

Discrimination in the UK: Leavers and Remainers

Abstract

We experimentally investigate in-group favouritism and out-group discrimination in the United Kingdom. We focus on the two groups formed following the 2016 EU Referendum (Leavers and Remainers). This paper uses social dilemma games: prisoner's dilemma, dictator game, trust game, public goods game and modified public goods game. We let individuals interact with in-group, out-group and individuals described as the national identity (British) to disentangle discriminatory behaviour. It also provides policy implications to identify potential anti-discrimination policies by emphasising national identity. The study also sets out to foster greater cooperation and test whether these groups can be further divided using social norm-based nudges. We find evidence of in-group favouritism and out-group discrimination. Beliefs about others' cooperativeness drive their contributions. Thirdly, in group settings - the Public Goods Games, the positive social norm-based nudge increases contributions from Leavers to Remainers, thus reducing discriminatory behaviour. However, it reduces prosocial behaviours intra-group for Remainers.

Keywords: Discrimination; Brexit; Identity; Leavers; Remainers; Polarisation; Norm-based nudge

3.1 Introduction

Political polarisation has been observed as increasing across the globe from France to Venezuela (McCoy et al. 2018, Dimant et al. 2020, Azmanova 2011, Mounk 2018). This political polarisation can threaten democratic practices, cooperativeness, altruism and trust (McCoy et al. 2018, Mounk 2018). Previous economic experimental research has shown that discriminatory behaviour can be dependent on characteristics such as ethnic groups (Ahmed 2010, Grosskopf & Pearce 2017, Chmura et al. 2016), gender (Eckel & Grossman 2001) and political ideology (Grosskopf & Pearce 2017, Abbink & Harris 2012). *In-group favouritism* is defined as the favourable treatment of a person with a shared identity. *Out-group discrimination* is the adverse treatment of a person with a different or opposing identity. Discrimination occurs along two main channels: taste-based discrimination, where agents gain utility from favouring their in-group above other groups (Becker 1971), or statistical discrimination, where agents have incomplete information and, therefore, rely on stereotyping (Arrow 1972). Therefore, discriminatory behaviour can arise from social preferences and beliefs (Güth et al. 1997, Guala et al. 2013, Ockenfels & Werner 2014). A recent identity group was formed following the 2016 EU Referendum in the United Kingdom on whether to leave or remain in the European Union - these two identities have now been coined *Leavers* and *Remainers*. The high salience referendum can be seen as a significant turning point for the United Kingdom. This paper examines how these newly formed naturally occurring groups of *Leavers* and *Remainers* impact behaviour and beliefs between these groups using a controlled online experiment. We also implement social norm nudge to shift discriminatory behaviour.

On the 23rd of June 2016, 51.9% of the UK electorate voted to leave the European Union (The Electoral Commission 2017). The decision to leave the EU arose from a slight majority of less than four percentage points (Ashcroft 2016). The referendum was proposed with no minimum qualified majority set and without a minimum turnout required (Bogdanor 2017). The rhetoric surrounding the referendum polarised citizens by emphasising the external other (Virdee & McGeever 2018, Levy et al. 2016). The 2016 referendum on EU membership questioned national identity. In addition, it has been argued that the 2016 referendum was used to show dissatisfaction with the government (Hobolt 2016) and that the underlying narrative behind the Leave Campaign was largely focused on the idea of `"Britishness"` (Virdee & McGeever 2018)[1]. Specifically, there are feelings of being left from globalisation, liberalism, and economic opportunities (Ashcroft 2016, Lawton & Ackrill 2016). Although very few of the British public had considered the role of the EU prior to the referendum,

[1]Notably, the division does not clearly align with any partisan group (Hobolt et al. 2021, Hobolt 2016).

the referendum soon proved to be a very dividing subject (Fetzer 2019, Hobolt et al. 2021). Recent studies have shown growing polarisation in the UK following the EU referendum (Virdee & McGeever 2018, Fetzer 2019). Notably, the formation of distinctive identities between Leavers and Remainers can be seen in reports; for instance, 38% of leavers would feel upset about the children marrying a Remainer and vice versa (Dalton 2019).

The referendum outcome can be seen as illustrative of the internal polarisation within the country, signalling a binary divide in which some feel they are winners (content with institutions and globalisation). In contrast, others feel like losers (discontent with institutions and globalisation) (Azmanova 2011, Fetzer 2019, Teney et al. 2013)—a similar conceptualisation of polarisation as in chapter 2. Briefly, citizens in nations experience polarisation when there is a movement towards extremes of the spectrum without any middle ground (Azmanova 2011, LeBas & Munemo 2019, Macy et al. 2019). Moreover, this study considers that polarisation has two dimensions, an external aspect and an internal aspect. This shared identity leads to raised uniting, which can further polarisation (Sunstein 2019*a*). Previous literature has observed that group polarisation can arise when the in-group norm is perceived as more extreme than the individual norm (Bonomi et al. 2021, Mill & Morgan 2021). Hence, in this study, our identification of polarisation specifically is two-fold. We look at self-reported alignment to the identities to identify movement to the extremes. Secondly, we compare differences between individual and group norms.

Polarisation rests on mechanisms of the theoretical frameworks of social identity. The economic and psychological literature has studied the impact of group membership on behaviours (Akerlof & Kranton 2000, 2005, Tajfel et al. 1971). Tajfel et al. (1971) observe that any association with an identity, whether a salient naturally-occurring identity or an artificial newly formed identity, impacts behaviours. Experimental research reveals that people favour their in-groups and discriminate against the out-group. Experimenters have found that naturally-occurring identities such as partisanship alters non-political behaviour between groups based on their political affiliations (Ben-Ner et al. 2009, Mill & Morgan 2021, Carlin & Love 2013). In our study, the voting decision is non-observable; therefore, in our experiment, voters must self-identify. The existing literature has relied on self-reported measures of identity characteristics that may not be observable, such as self-reported voting decision in the context of the United States (Mill & Morgan 2021), self-reported levels of religious affiliation (Eckel & Grossman 2004), self-identified as either Right or Left-wing (Thomsson & Vostroknutov 2017), in aliis. Therefore, we posit that the decision to vote in the 2016 EU Referendum formed a group identity, though unobservable, but still created a

sense of collective identity and formed group membership.

The study by Abbink & Harris (2012) used an allocator game and dictator game to measure discrimination between political groups in Thailand amongst students. The authors find that allocators strongly favour their in-group and discriminate against rival party supporters. The authors identify the difference between in-group favouritism and out-group discrimination. They find evidence of both in-group favouritism and out-group discrimination. Likewise, Ben-Ner et al. (2009) finds that both in terms of giving in the dictator game and selecting people as a part of a team, identities impact behaviours, including political views[2]. For instance, Carlin & Love (2013) finds that partisanship affects trust between Democrats, Republicans and Independents players, with greater trust along shared party lines and discrimination across party lines. Mill & Morgan (2021) find that political affiliation in America between Trump and Clinton supports leads to dysfunctional behaviour.

Similarly, Abbink & Harris (2019) notes that to effectively identify discriminatory behaviour between rival groups, this needs to be considered with reference to a non-rival group. Perhaps an obvious non-rival group is an identity that is shared by rival groups, such as the national identity. For instance, Robinson (2016) finds that strong national identification alters discriminatory behaviour between ethnic groups. Our study focuses on two distinct (rival) groups - Leavers and Remainers. We also include the non-rival shared national identity group of the British identity. Here we aim to analyse the impact of Brexit on behaviours and understand to what extent being a Leaver or Remainer will impact economic decision-making in non-political contexts. Specifically, we investigate if participants behave differently between the in-group, out-group and neutral or non-rival group.

The underlying cause of discrimination is partly due to acceptance of the status quo. Much of the literature generally emphasises that an individual's behaviour is affected mainly by incentives, norms and customs (Elster 1989, Green 2012, Krupka & Weber 2013, Ostrom 1998, Barr et al. 2018, Bénabou & Tirole 2011*a*, Bowles & Polania-Reyes 2012). Research has shown that social norms dictate behaviour in the private, social and political spheres of life (Barr et al. 2018, Harris et al. 2015, Hornung et al. 2019, Krupka & Weber 2013, Ostrom 1998). The social nature of norms means that they are values that a certain identity group shares (Krupka & Weber 2013, Cialdini et al. 1991, Farrow et al. 2017). These norms are self-enforced in social contexts. Hence, different settings, such as diverse types of individuals, locations, cultures and times, will affect the social norms drawn upon (Ostrom 1998). There-

[2]We refer to similar literature as chapter 2, however, here, we focus more on experiments that use politically polarised participants.

fore, as individuals engage in voting decisions, their decisions may be based on conformity to a particular identity (Bonomi et al. 2021). Notably, Apffelstaedt et al. (2022) finds that perceptions of the socially appropriate action of sharing income with poorer individuals shift once an elected official. These findings suggest that social norms are able to be moulded and altered by the political environment. Nevertheless, norms are not constantly prevalent in individual decision-making. Therefore experimentalists need to ensure that these norms are brought into the economic decision-making process in experiments (Krupka & Weber 2013). Thus, to explicitly identify the role social norms and perceptions of majority Leavers and Remainers play in decision-making, participants undertook a social norms elicitation task similar to that established by Krupka & Weber (2013).

Social norm-based interventions that utilise the prevalence of social norms have become increasingly popular in experimental and policy research (Mol et al. 2020, Bicchieri 2015, te Velde & Louis 2022, Bicchieri & Dimant 2019, Holzmeister et al. 2022, Allcott 2011). Norm-based interventions encourage individuals to act according to modal behaviour by anchoring participants to a specific reference point Bruns et al. (2018). Behaviourally informed policies are mainly focused around behavioural nudges - influencing an individuals behaviour to a desirable outcome without damaging their freedom of choice (Thaler & Sunstein 2008), information techniques and more recently, behavioural sludges - making an action more effort or cumbersome to deter them from doing it (Lorenz-spreen et al. 2020) to achieve desired policy outcomes (Thaler 2018). Behavioural policy tools, such as these, vary in application from women in the workforce, energy consumption, alcohol consumption and gender roles (Abrahamse et al. 2005, Bicchieri 2015, Espinosa & Stoop 2021). Behaviourally informed interventions constitute an updated methodological toolkit available for policies and initiatives. Nudges have become a policy tool used by many governments (Hansen & Jespersen 2013, Thaler 2018, Thaler & Sunstein 2008, Nagatsu 2015).

In this paper, we also consider *how* discriminatory behaviour can be alleviated and conversely further divided. Bicchieri (2015) describe how norms begin to alter individuals' behaviour when considering gender roles. People generally form beliefs about others to explain divisions between groups. More specifically (injunctive) social norms are those that 'prescribe what one should do or should not do'. We focus on injunctive norms - what others find appropriate behaviour (Cialdini et al. 1991, Farrow et al. 2017, Cialdini et al. 2006, Heinicke et al. 2022). In a similar study, Hanel & Wolf (2020) provide information on the similarities between Leavers and Remainers. They find that information on the similarities between Leavers and Remainers can help to mitigate the void between these groups. In our paper, we lean more

heavily on injunctive norms, both prescriptive and proscriptive.

Our experiments consist of three main stages. Firstly, at the beginning of the experiments, participants were asked what decision they made in the 2016 EU Referendum. In the second stage of the study, we use a norm-based intervention to reduce discriminatory behaviour between these groups. We use a similar social norm-based nudge intervention used in chapter 2 to understand how specific policy tools such as nudges can have various or different responses in different contexts (Howlett et al. 2020, Howlett & Leong 2022, Thaler & Sunstein 2008, Graf 2019, Nagatsu 2015). As in chapter 2, we exogenously manipulate the prevailing norms before undertaking the games, using two social norms: *Heal*, a prescriptive social norm to foster greater prosocial outcomes between the two opinion-based groups and *Divide*, a proscriptive social norm to separate the two groups further. Following the treatments, participants play non-strategic games, which arise when individuals are not motivated by future material rewards, as there is no future interaction between participants[3]. We use a variety of within-subject experimental games and reveal the receiver's identity as either: a Leave voter, Remain voter and British citizen, using well-documented experiments in the literature - the prisoner's dilemma, dictator game, trust game and public goods game. Discriminatory is measured by differences in behaviour depending on the revealed identity of the other player and in-group favouritism as preferential treatment to in-group members (Güth et al. 1997, Guala et al. 2013, Ockenfels & Werner 2014). This paper also considers in-group and out-group actions in our social norms elicitation task. We compare an individual's perceptions of certain scenarios and the perceptions of the in-groups and out-groups actions in different scenarios.

We find evidence of out-group discrimination and in-group favouritism amongst Leavers and Remainers. In addition, the other player's beliefs significantly influence behaviours across all games. We find mixed results of the social norm-based interventions—both treatments *bridge the gap* between Leavers and Remainers in the Public Goods Game when the sender is a Leaver. However, the interventions backfire, leading to reduced in-group cohesion for Remainers. This paper also observes that norm-based interventions significantly alter norm evaluations. Specifically, the randomly assigned participants in the *Divide* treatment significantly altered their evaluations of social norms following the treatment.

This paper contributes to the literature on discriminatory behaviour between naturally occurring groups recently formed and chosen by individuals. This paper investigates the degree

[3]The extant literature has investigated whether strategic games are strategic and if significant differences in behaviour between the two game settings can be observed (see for example (Reuben & Suetens 2012)

to which the 2016 EU referendum vote created distinct groups and if this political polarisation leads to discriminatory behaviour. We use incentivized behavioural games using an online sample of participants recruited through Prolific. We also measure the extent of the divide using self-reported identity salience as a quantitative measure of polarisation. The paper also extends the behavioural governance literature by examining social-norm-based nudges' effectiveness in shifting behaviours between groups. This thesis brings together the gap whereby social categories are harder to observe in nature yet are dominant. Also, the context is heightened by having small diversity among these social groups.

The remainder of this paper is organised as follows. In section 3.2, we identify the theoretical framework and hypotheses. In section 3.3, we present our experimental design and procedure. Section 3.4 reports our results on discriminatory behaviour between Leavers and Remainers. Finally, Section 3.5 discusses the implications of our results and concludes the paper.

3.2 Theoretical Framework and Hypothesis

We replicate the theoretical underpinning used in chapter 2. However, here we use a simple model based on political identity. In sum, the concept of group polarisation is closely associated with social identity formation and inter-group conflict. Various theories have been identified to explain in-group favouritism and out-group discrimination. Social identity theory (SIT) posits that people either maximise in-group outcomes or differences between in-group and out-group outcomes, even in minimal or artificial groups (Tajfel et al. 1971). Agents identify themselves based on mere self-categorisation into different groups separate from their personal or individual traits. This categorisation has an emotional response which leads to divergent behaviours between groups (Balliet et al. 2014). Akerlof & Kranton (2000, 2005) stress the importance of identity in influencing economic behaviour. The authors posit that economic agents gain utility from behaving appropriately with respect to their social group. Identity with a particular group will impact behaviour with in-group members and out-group members (Chmura et al. 2005). This identity is internal to agents, and there may be the threat of sanction if one deviates from the group's norms. Even when one is by themselves, the identity they proscribe themselves, and the associated social norms guide them to a certain behaviour, even at the expense of individual interest (Elster 1989). Hence, to avoid sanctioning and maintain positive self-esteem, agents conform to the social identity, even at the expense of economic efficiency, often viewing their group as morally right, whilst the out-group is the opposite.

We use a simple model of utility maximisation that closely follows Akerlof & Kranton (2000, 2005), Mill & Morgan (2021), as well as the role of norm-compliance and identity mirroring Krupka & Weber (2013). An individual $i's$ utility $U(a_i)$ depends on both their actions and the degree to which their actions comply with the prevailing norms $N(a_i)$:

$$U(a_i) = V(\pi(a_i) + \gamma N(a_i) + (\eta)(\pi_i(a) - \pi_j(a)) \tag{3.1}$$

The component V represents the value an individual places on the material payoff of action a. The function increases in $\gamma(a_i)$, $\gamma \geq 0$, as it increases, an individual will get higher utility from complying with the norm - in other words, doing the socially appropriate action. With $N(a_i)$ being a norm function. If people were completely concerned about economic efficiency $\gamma = 0$, they would always choose the option that maximised payoff. Deviation from the social norm yields disutility. This fear of deviation or disapproval of others can generate a "psychological cost" (Fehr & Schurtenberger 2018). The "social multiplier" (Bénabou & Tirole 2011*b*, p.6), is likely to drive a shift in behaviour in line with the social norm, which we exogenously manipulate in the *Heal* and *Divide treatments*. Participants with a higher γ will be more inclined to comply with cooperative social norms even in one-shot games. However, lower levels of γ, suggest that these participants are on the threshold and are more likely to deviate from the social norm.

Let $\eta = 0$ if participants share the same political identity as their opponent, and $\eta = 1$ otherwise. The utility of individual i is increasing in π_i. It is evident that the decreasing utility in $\pi_j(a)$ is larger if the participants do not share the same political identity as their opponent. Thus, we hypothesise that polarisation will be more significant if they are paired with an opposing voter compared to a coinciding voter, i.e. political identity influences antisocial behaviour. Informing agents of the prevailing social norms should bring to the foreground the group benefits and costs relative to the intrinsic benefits and costs, increasing the influence of conformity (Sunstein 2019*a*). Therefore, agents yield higher utility from maintaining positive self-esteem by conforming to group behaviour. In addition, polarisation leads to stronger identification with the in-group (movement to the poles), further impacting discriminatory behaviour.

Hypothesis 5 *Participants behave more discriminatory when paired with the opposing voter.*

Hypothesis 6 *The more strongly individuals identify with their ethnicity (the more polarised), the greater the discriminatory behaviour.*

Hypothesis 7 *Following the* Heal *treatment, participants will be more prosocial towards out-group members.*

Hypothesis 8 *Following the* Divide *treatment, participants will be less prosocial toward out-group members.*

3.3 Experimental Method

In the following, we present the experimental design, the procedure and other measures of interest. To test our hypotheses, we used a similar experimental method as used in chapter 2. However, in this paper, we included a variation of one of the experimental games (see subsection 5.3.7), differing vignettes in the social norms elicitation task (see subsection 3.3.7), and different norm based-intervention (see hd and polarisation measures (see subsection 3.3.8). In addition, the experiments were conducted online.

3.3.1 Heal and Divide

At the beginning of the experiment, participants are randomly assigned to the treatment conditions. Participants received information about the prevailing social norms. Participants in the first treatment group are given feedback on previous respondents with shared social norms from the in-group and out-group, labelled the *HEAL* treatment. In the second treatment group, participants are given feedback on different and diverging social norms between the in-group and out-group, labelled the *Divide* treatment. In the control treatment, no information is provided [4]. Participants will be given the social norms information through charts that show the proportion of participants with shared views of social norms of each scenario at the beginning of the session see Figure 3.1 and Figure 3.2, and as a baseline group, we include a *Control* that receive no information. After being placed into one of the three treatment groups, participants were given the same games, social norms elicitation task and post-experiment questionnaire. The random allocation of the intervention occurs before the games are played so that social norms are primed. This brings norms to the forefront and increases norm compliance.

[4]Here, we deviated from the previous paper by using a different norm scenario. The scenario of vandalism was chosen due to the increased reports of hate crime and vandalism following the referendum (Virdee & McGeever 2018, Komaromi 2016, Albornoz et al. 2021). In addition, we intentionally chose a loaded norm to induce behavioural change (Abbink & Hennig-Schmidt 2006)

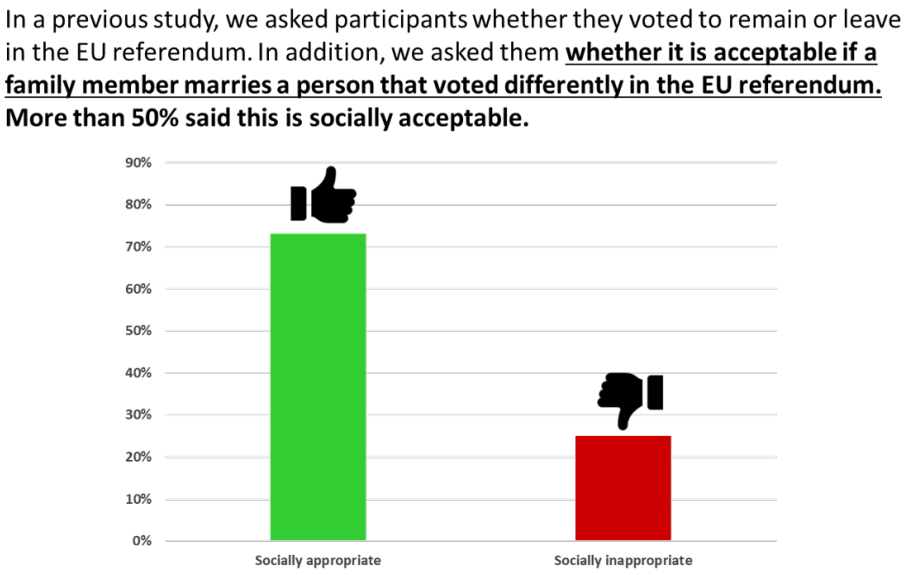


Figure 3.1: Social Norm nudge (HEAL)

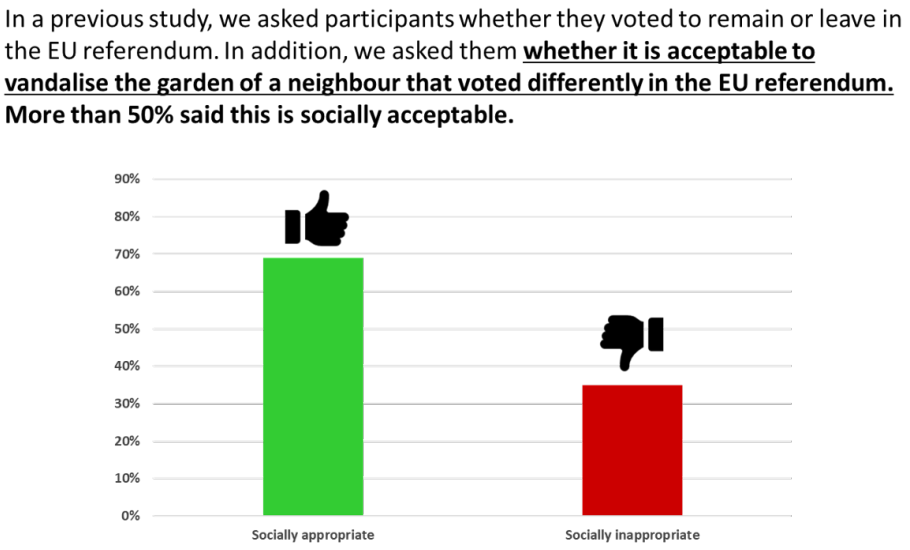


Figure 3.2: Social Norm nudge (DIVIDE)

3.3.2 Prisoner's Dilemma (PD)

Table 3.1 presents the payoff matrix for the one-shot prisoner's dilemma. Players can choose between two options (strategies) - cooperate or defect. Each participant is aware of how much they could gain, but this depends on the other participant's choice. Participants who choose an outcome with the least difference indicate a preference to cooperate (C, C); this is

the Pareto-dominant strategy since there is no other option that gives both players a higher payoff. However, preferences to defect are indicated by participants choosing the outcome with the highest individual payoff, even if there is a significant difference between payoffs between participants (Cardenas & Carpenter 2008). Defecting is the dominant strategy (D, D).

Participants were given neutral language and shown a table with Red and Blue cards. The cooperative outcome is (B,B), and the purely self-interested outcome is (R,R) see Figure A.1. Participants are tasked with making incentivized decisions between an in-group, out-group and mixed-group (British) receiver; this indicates if preferences toward cooperation change depending on the social identity of the partner. Differences in cooperation between receivers measures taste-based discrimination. We also elicit participants' beliefs regarding their beliefs about the decisions of others. Therefore, we can identify statistical discrimination [5].

Table 3.1: Payoff Matrix (Prisoner's Dilemma)

		Player 2	
		Defect	*Cooperate*
Player 1	*Defect*	(1,1)	(10,0)
	Cooperate	(0,10)	(5,5)

You Interact	Other Interacts	You Win	Opponent Wins
		5	5
		0	10
		1	1
		10	0

Figure 3.3: Prisoner's Dilemma in the experiment in points

[5]We elicit beliefs after the one-shot games. We then match their beliefs ex post (De Geest & Stranlund 2019, Schotter & Trevino 2014, Gächter & Renner 2010). The literature has shown that beliefs about others' behaviour can affect instances of cooperation and prosocial behaviour (Fischbacher & Gächter 2010, Gächter & Renner 2010, Guala et al. 2013). As participants may expect less cooperative and prosocial behaviour from in-group members (Ferraro & Cummings 2007). Gächter & Renner (2010) noted that beliefs were more accurate when incentivized; for this reason, we incentivized beliefs.

3.3.3 Dictator Game (DG)

The dictator game by Roth & Erev (1995) tests if individuals act out of pure self-interest. Player A is endowed with 10 points for each receiver (a Leave, Remain and British player). If Player A allocates an amount to Player B, Player A gets to keep the remaining amount. A fair outcome would be to share the endowment equally between the two participants. The amount of points the sender decides to give to the other player measures taste-based discrimination (Lane 2016, Grosskopf & Pearce 2017). Participants' beliefs regarding the decisions of others are used to identify statistical discrimination.

3.3.4 Trust Game (TG)

The trust game by Berg et al. (1995) identifies the degree of trust and reciprocity between participants. Player A is endowed with 10 points for each receiver (a Leave, Remain and British player). Player A is tasked with deciding how much of their endowment (x) if any, to send to player B. Any amount player A does not send, they can keep. The amount they send to player B is multiplied by 3. Player B receives the amount sent by participant A ($3x$). Player B then decides how much to return (y) to player A, with $y \leq 3x$.

Final payoffs are $P_A(x,y) = E - x + y$ for player A, and $P_B(x,y) = 3x - y$ for player B.

The amount transferred (x) measures player A's trust. Any positive amount sent from Player A indicates that they trust either the in-group, out-group or mixed-group to reciprocate any positive contribution. However, Player A is vulnerable as Player B may choose to keep all the positive contributions for themselves. Therefore, the trust game measures taste-based and statistical discrimination, as senders consider the potential for reciprocity. Again, we also elicit participants' beliefs regarding the decisions of others to identify statistical discrimination or stereotyping.

3.3.5 Public Goods Game (PGG)

Participants play the game with three other group members in the public goods game (PGG) by Gächter & Fehr (1999). Participants make contributions to the common pot. The total contribution is multiplied by a factor α ($1 < \alpha < n$). The contributions will then be equally re-distributed between all four participants: the PGG tests cooperativeness and the degree

of prosocial behaviour.

The design uses a standard linear public goods game played by four n participants (including themselves). Each participant is endowed with 10 points for each receiver and is asked to decide how much to keep or contribute to a public good. The sum of the group's payoff is multiplied by 1.5, which is shared equally between the four participants. The PGG occurred with homogeneous groups or politically mixed groups. More specifically, participants are tasked with making incentivized decisions between in-group $(n-1)$, out-group $(n-1)$ and mixed-group $(n-1)$ receivers. The final payoff for the game for each participant is given by:

$$\pi_i = 10 - g_i + \frac{1.5}{4} \sum_{j=1}^{4} g_j \tag{3.2}$$

The public goods game has two equilibria, a Nash equilibrium and a Pareto optimal outcome. The Nash equilibrium is the selfish payoff maximising dominant strategy where participants free-ride and do not contribute to the common pot, keeping their endowment but benefiting from the divided shares. The Pareto (or socially) optimal equilibrium occurs when you cannot make any participants better off. In the PGG, this occurs when there is full cooperation, put differently, when all players contribute their entire endowment to the common pot. The Pareto optimal outcome suffers from being a week equilibrium, with the alternative non-cooperative equilibrium being stronger. In repeated games, if someone from the group deviates from the socially optimal outcome due to free-riders, more and more participants will free-ride over time. At the alternative equilibrium, if an individual deviates and begins to contribute more than zero, contributions will quickly go back to zero. In this study, we use the one-shot public goods game to examine taste-based discrimination in one-shot games difference in contribution level between in-group and out-group matching. We also elicit the beliefs, as players decide how much they believe others would send to them.

3.3.6 Modified Public Goods Game (mPGG)

Figure 3.5 presents the modified public goods game to identify if group size matters and if a critical mass significantly alters altruism and cooperation in a similar decision structure. Participants will play the game with one group member in the modified public goods game (mPGG). The mPGG tests cooperativeness and the degree of prosocial behaviour. Each participant is endowed with 10 points for each receiver and is asked to decide how much

to keep or contribute to a public good. The sum of the group's payoff is multiplied by 1.5, which is shared equally between the four participants. Participants are tasked with making incentivized decisions between in-group, out-group and mixed-group players. The Pareto optimal outcome and Nash equilibrium are the same as the PGG. We also elicit the beliefs, as players decide how much they believe others would send to them, see Figure 3.5. The final payoff for the game for each participant is given by:

$$\pi_i = 10 - g_i + \frac{1.5}{2} \sum_{j=1}^{2} g_j \tag{3.3}$$

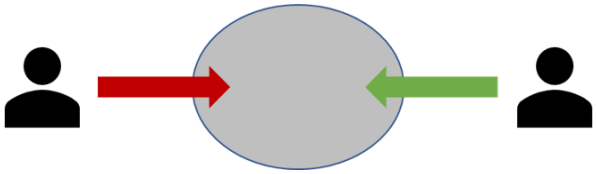


Figure 3.4: Modified Public Goods Game - Sender

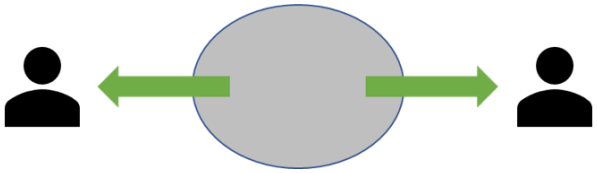


Figure 3.5: Modified Public Goods Game - Beliefs

3.3.7 Measuring social norms - Social Norms Elicitation Task

The elicitation of participants' social norms is an adaption of the method by Krupka & Weber (2013). This paper deviates from Krupka & Weber (2013) by using real-world scenarios rather than the allocator game. Participants are tasked with evaluating the social appropriateness of real-world vignettes. This removes the abstract nature of game-based situations, enabling more realistic ideas of prevailing social norms. Additionally, participants are asked to evaluate each scenario three times, first considering their perceptions of the social appropriateness of the provided action, then what most people that belong to their in-group and what most people that belong to a different identity group consider socially appropriate (Akerlof & Kranton 2005).

Participants' beliefs about social norms are incentivized, as they are aware that their decisions will be matched with the modal response of Leave voters and Remain voters. The incentivization of the social norms task is vital as norms are perceived collective behaviours (Akerlof & Kranton 2005, Elster 1989, Green 2012, Ostrom 1998). In this study, participants are also asked to evaluate their perceptions before the modal Leaver and Remainer perceptions of the norms. Participants are asked to evaluate the scenarios to provide insight into the difference between personal convictions and social norms. Participants evaluate the following scenarios:

- *Vandalism 1* - Suppose someone vandalises the garden of a neighbour after becoming aware that they voted to leave the EU.

- *Vandalism 2* - Suppose someone vandalises the garden of a neighbour after becoming aware that they voted to Remain in the EU.

- *Vandalism 3* - Suppose someone vandalises the garden of a neighbour after becoming aware that they Did not vote in the EU Referendum

- *Family Scenario* - Suppose someone in your family wants to get married to someone that voted differently to them in the 2016 Referendum.

- *Briber Scenario 1* - You see a manager bribing a public official to get a contract. How appropriate is it to report the manager.

- *Bribe Scenario 2* - A citizen is made aware of this bribe. How appropriate is it to punish the manager.

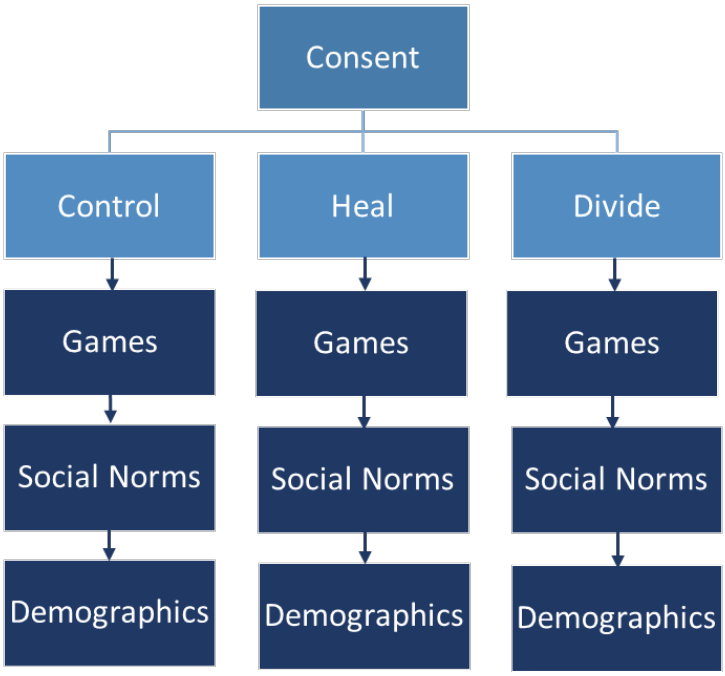


Figure 3.6: Experimental Summary

3.3.8 Procedure

The experiments were conducted in 2021. However, the referendum occurred several years prior to the experiment. The fallout of recently leaving the EU remained a pertinent topic at the time of the experiments. While the referendum occurred in the past, the consequential effects of this pivotal event continue to shape public discourse, policy considerations and individual perceptions. Therefore, when the experiments were conducted, Brexit remained contextually relevant. The experiments were conducted online using the Prolific recruitment service. The experiments were programmed using the online survey software Qualtrics (Qualtrics 2020). At the beginning of the session, participants received instructions. The instructions used were neutral and aimed not to detail the experiment's objectives. See Appendix B. They were informed that they would receive a fixed participation fee (£2.50) and an additional payoff which would differ for each participant. To determine their payoff, participants were truthfully told that one of their choices would be randomly selected at the end of the experiment. All participants began with the games, then the social norm elicitation task [6]. Participants had to play with the following receivers in all games: Leavers, Remainers

[6]The social dilemma games are randomised to prevent any ordering effects. Note that the social norm elicitation task occurred after the experimental games, although D' Adda et al. (2016) finds limited evidence of order effects of the norm elicitation task. We use a within-subject design to control for idiosyncrasies in

and British. Please note that throughout the survey, the UK citizen participant will be with a 25% probability a participant from (England, Scotland, Wales, and Northern Ireland). Participants did not receive any feedback from the games undertaken. All participants in the sample passed the attention checks. See Figure 3.6 for summary of experiments.

In addition, to shed light on their subjective views on their identities, we asked participants questions about their perceptions in the post-experiment questionnaire. First, we asked participants how much they trusted in-group and out-group members when interacting. We also included the government as someone whom they also interact with. Second, we asked participants to evaluate their aversion to risk in general, then more specifically in financial matters, interacting with the government and their health. Given that during and following the aftermath of the 2016 EU Referendum, the prevailing discourse surrounding Brexit at times focused on the risks of leaving or staying in the European Union on the nation, economy and people (Levy et al. 2016, Virdee & McGeever 2018, Hobolt 2016, Fetzer 2019). Indeed, risk aversion has been found to explain differences in behaviours between genders (Iriberri & Rey-Biel 2021). Also, risk aversion may confound altruistic behaviours (Eckel & Grossman 2008). Therefore, we are also interested in controlling for risk-seeking behaviour that may relate to voting decision and propensity for pro-social behaviour.

Each question in the self-reported polarisation measures is to be answered on a 5-point Lickert scale. The self-reported polarisation questions were posed, such as `"to what extent do they align with their in-group identity"`. Fourthly, we asked participants to evaluate how proud they are to be their in-group identity. The question is coded to run from 1 to 4, 1 being `"not proud at all"` to 4 being `"very proud"`. Lastly, we asked participants if they felt like a winner following the EU referendum, which is equal to 0 if you feel like a `"loser"`, equals one if you feel like a `"winner"`. See Appendix B for each question.

Additionally, to ensure that the heterogeneity between these two homogenous groups of interest is not neglected, data such as sex, education, ethnicity, employment status, household size, and marital status, as well as information such as age and income variables, are collected to use as controls in the analysis of the research.

participants (D' Adda et al. 2016)

3.4 Results - Discriminatory behaviour and Polarisation

We present our results in multiple steps. We begin by analysing discriminatory behaviour dependent on the identity of the receiver. We then identify individual and majoritarian evaluations following the social norm elicitation tasks. Lastly, the study investigates the impact of the social norm-based intervention on behaviour and norm compliance.

Table 3.2: Summary Statistics

	All	**Control**	**Heal**	**Divide**
Identity				
Leave voter (%)	47.33	44.51	49.38	48.26
Remain voter (%)	52.67	55.49	50.63	51.74
Socio-demographic characteristics				
Age				
25-34	36.63	35.26	35.63	38.95
35-44	25.54	26.01	23.75	26.84
45-54	19.41	24.28	18.75	14.21
Female (%)	72.67	67.63	76.25	74.42
University Degree (%)	68.71	70.52	69.38	66.28
Single (never married) (%)	48.12	50.87	48.13	45.35
Household size	2.92	2.89	3.03	2.85
Nationality (%)				
English	86.96	86.56	88.59	85.79
Northern Irish	7.79	1.08	1.09	3.16
Scottish	8.04	10.22	5.98	7.89
Welsh	3.21	2.15	4.35	3.16
Ethnicity (%)				
White	90.89	93.01	89.13	90.53
Risk	4.18	4.03	4.54	3.99
Polarisation measures				
Proud of voting decision	83.56	84.97	83.75	81.98
Winner (%)	33.47	35.26	31.88	33.14
Highly aligned to identity (%)	61.19	61.27	61.88	60.47
Revote (%)	74.65	77.46	74.38	72.09
Observations	505	173	160	172

Notes: Baseline summary statistics of participant characteristics in the experiment

Table 3.2 summarises the characteristics of participants in the experiment. Across the intervention treatment variations, the percentage of Leave voters and Remain voters are relatively similar. Other socio-demographic factors, as well as the polarisation measures, are relatively

constant across treatments. This supports the randomisation of participants into intervention treatment groups. However, we obtained an unbalanced sample of participants containing a relatively high proportion of White English Females. Therefore, we ensure that we include gender, ethnicity and race as covariates in regression models to control for the potential confounding effects of these factors in behaviours, to address the possibility of results being skewed by the sample composition.

Consistent with Alabrese et al. (2019), we find that specific characteristics are significantly different between Leavers and Remainers. We find education is significantly different between Leavers and Remainers (Wilcoxon rank-sum test p=0.000, Kruskal-Wallis test p=0.000). Age is also significantly different between the two identities (Wilcoxon rank-sum test p=0.000, Kruskal-Wallis test p=0.000). As well as gender (Wilcoxon rank-sum test p=0.001, and Kruskal-Wallis p=0.001), respectively. Similarly, we also find that nationality and ethnicity significantly differ between Leave and Remain voters (Wilcoxon rank-sum test p=0.000 and p=0.027, respectively). However, our experiments do not find employment status significantly different between Leavers and Remainers.

We examine the relationship between the self-reported polarisation measures and the participants voting decisions. There is a statistically significant difference in the *proudness* of voting decisions between Leavers and Remainers (Wilcoxon rank-sum test p =0.000). We observe a similar statistical difference regarding feeling like a *winner* (Wilcoxon rank-sum test p =0.000). How highly a participant *aligns* with their voting decision is statistically significantly different between Leavers and Remainers (Wilcoxon rank-sum test p =0.000). We find a similar relationship regarding *revoting* (Wilcoxon rank-sum test p =0.000).

3.4.1 Prisoner's Dilemma (PD)

Turning now to consider the decision of the prisoner's dilemma. In our prisoner's dilemma, zero equals Defect, and one equals Cooperate. In Figure 3.7, we find that the beliefs of others, specifically, Leavers and UK players are less than actual transfers. However, when the receiver is a Remainer, beliefs are greater than transfers. Figure 3.8 shows the transfers by each identity relative to the mean. Participants who voted to Leave the EU coordinate more than the mean when a receiver is a fellow Leave voter and a UK receiver but less than the mean level of coordination when the receiver is a Remainer. When the participant (sender) is a Remain voter, they coordinate more than the mean level of coordination when coordinating with Remain (in-group) receivers but less than the mean for both Leave and

UK receivers.

To identify taste-based discrimination, we compare contributions between the two groups. We use Wilcoxon rank-sum tests to test differences in coordination between Leave and Remain voters in the prisoner's dilemma game. When coordinating with a Leaver, coordination is significantly different between Leave and Remain voters ($p = 0.000$). We find a similar result when coordinating with Remain voters ($p = 0.000$). We do not find a statistically significant difference in coordination between Leavers and Remainers when coordinating with a UK citizen (p =0.681 and p =0.357, respectively). Therefore, we find evidence of taste-based discrimination.

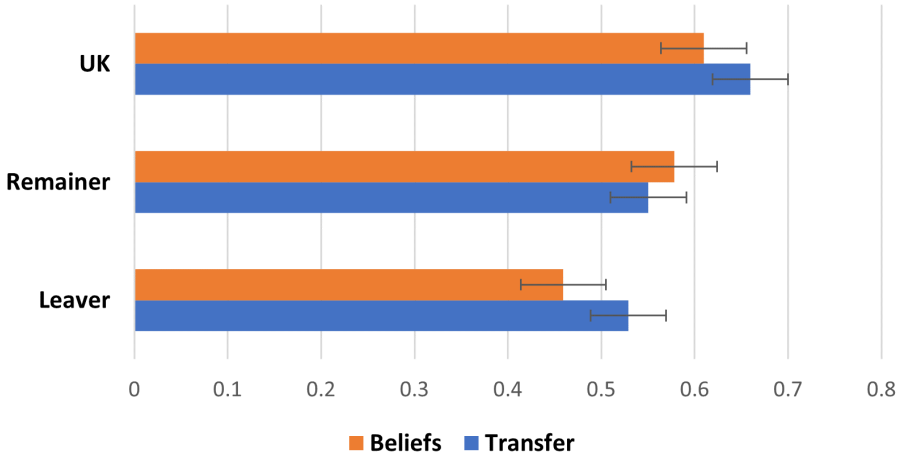


Figure 3.7: Prisoner's Dilemma

Note: The figure shows the percentage of participants choosing coordination (=1) and the beliefs of others choosing to coordinate. The orange bars present beliefs, and the blue bars presents the fraction of coordination for each receiver (UK, Remainer and Leaver). Error bars indicate 95% confidence intervals.

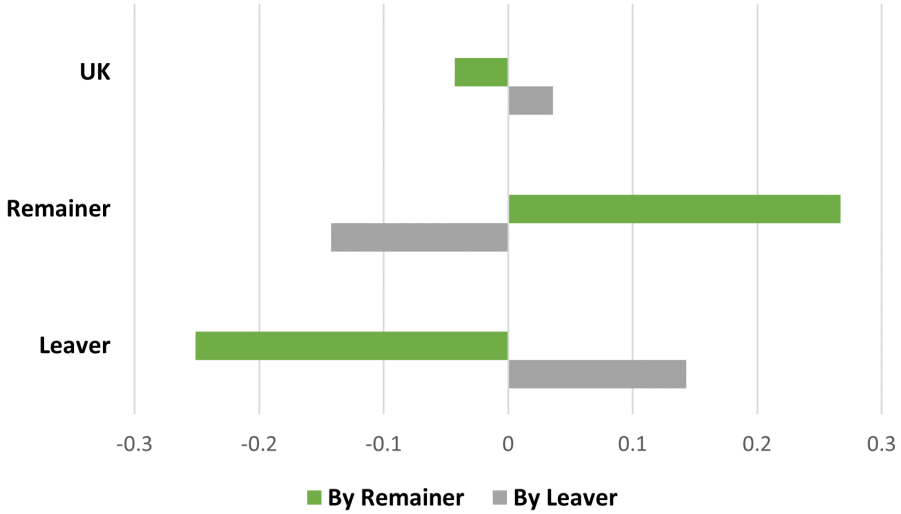


Figure 3.8: Prisoner's Dilemma (mean-centred)

Note: Zero line shows the mean benchmark. The orange bars show differences from mean transfers by Remain voters. The blue bars show transfers different from the mean by Leave voters.

Table 3.3 reports the results of Probit regression analysis of behaviour in the prisoner's dilemma on identity. The dependent variable is the likelihood of coordination zero equals Defect, and one equals Coordinate averaged for when the receiver is a Leaver (column (1) and (2)), Remainer (column (3) and (4)), and UK citizen (column (5) and (6)), we include socio-demographic controls for age, gender, marital status, employment status, household size, educational attainment, religious indicator, race and ethnicity. However, Leave voters are less likely to coordinate when the receiver is a Remainer.

Our model also includes beliefs reported during the experimental games. The results show that beliefs about the receiver significantly impact coordination, generally increasing the likelihood of coordination across all receivers. Polarisation is measured using variables identified in the post-experiment questionnaire. We find no evidence of self-reported polarisation measures impacting the likelihood of coordination in the prisoner's dilemma.

Table 3.3: Probit regression Prisoner's Dilemma by contribution

Dep. variable: Coordinate(=1, 0 otherwise)	Transfer Leaver		Transfer Remainer		Transfer UK	
	(1)	(2)	(3)	(4)	(5)	(6)
Leaver	0.428***	0.304	-0.509***	-0.478**	0.109	0.200
	(0.121)	(0.190)	(0.121)	(0.190)	(0.122)	(0.197)
Beliefs	0.911***	0.951***	0.951***	0.980***	1.082***	1.147***
	(0.122)	(0.128)	(0.122)	(0.125)	(0.122)	(0.128)
Polarisation						
Vote proud		-0.030		-0.001		-0.062
		(0.083)		(0.081)		(0.082)
Winner		0.071		-0.034		-0.301
		(0.176)		(0.168)		(0.186)
Align		0.134		-0.081		0.056
		(0.135)		(0.138)		(0.144)
Controls	NO	YES	NO	YES	NO	YES
Constant	-0.535***	5.425***	-0.166	-5.566***	-0.250**	1.019
	(0.090)	(0.696)	(0.115)	(0.704)	(0.106)	(0.910)
Observations	505	504	505	504	505	504

Note: Standard errors clustered at the subject level in parentheses.*** p<0.01,** p<0.05,* p<0.10.

3.4.2 Dictator Game (DG)

Figure 3.9 presents the transfers and beliefs in the dictator game. In the dictator game, beliefs are lower than actual transfers to all receivers. The average amount transferred to Leave receivers is 3.34 points from the 10 points endowment. Remain voters, on average, received greater transfers at 4.08 points. UK receivers received the greatest points on average at 4.20 points. Figure 3.10 presents amounts sent that diverge from the mean transfer by the

two identities. The figure shows that Leavers transfer more than the mean to fellow Leave voters. Conversely, Remainers transfer greater than the mean to fellow Remainers and UK players.

To test differences in contributions between Leave and Remain voters in the dictator game, we use Wilcoxon rank-sum tests. When the receiver is a Leave voter, there is a statistically significant difference in contributions between Leave and Remain participants (p =0.000). We find a similar result when the receiver is a Remain voter (p =0.000). When the receiver is from the United kingdom, we do not find a statistically significant difference in contributions (p =0.313).

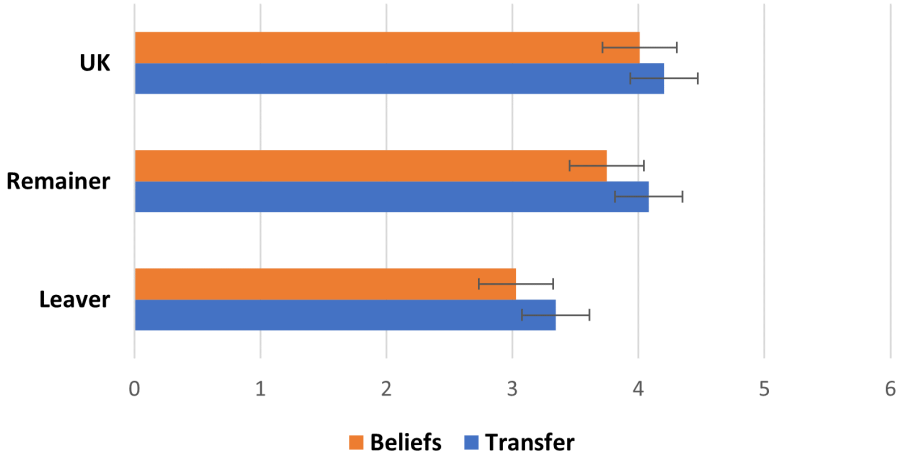


Figure 3.9: Dictator Game (Mean contributions and Mean Beliefs)

Note: The figure shows the mean average transfers and beliefs in the Dictator Game. The orange bars present beliefs, and the blue bars present the amount sent for each receiver (UK, Remainer and Leaver). Error bars indicate 95% confidence intervals.

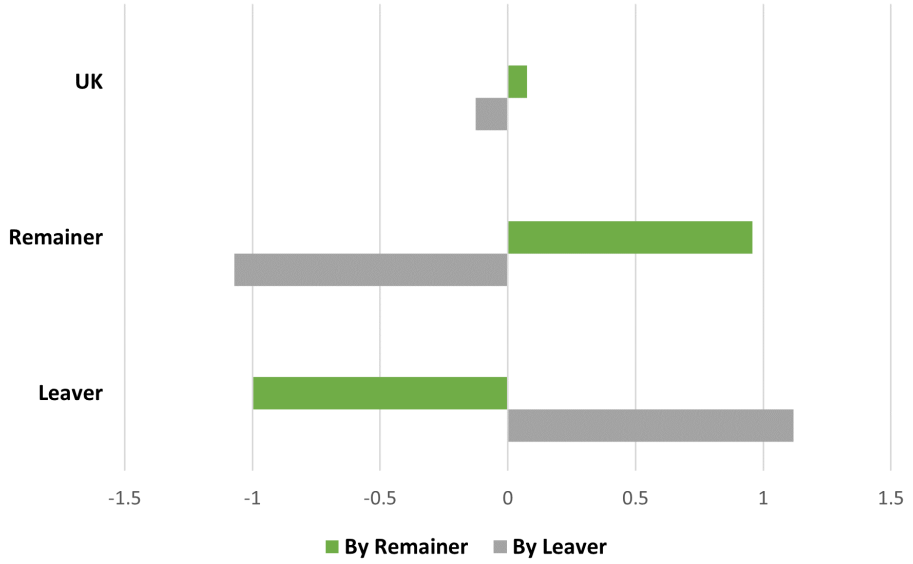


Figure 3.10: Dictator Game (mean-centred)

Note: Zero line shows the mean benchmark. The orange bars show differences from mean transfers by Remain voters. The blue bars show transfers different from the mean by Leave voters.

Table 3.4 reports the estimated OLS models of contributions in the dictator game controlling for age, income, sex, educational attainment, age, income, household size, ethnicity, race, marital status and risk-seeking. When the participant is a Leave voter, we observe less cooperation when the receiver is a Remainer; however, this is not robust to the inclusion of controls. Beliefs significantly impact the behaviours of all receivers. Therefore, perceptions of others significantly impact prosocial behaviour.

We use self-reported measures to examine how participants' behaviour is influenced by polarisation. We find that feeling like a *winner* following the Referendum reduces cooperation when the receiver is a Remainer. In addition, being highly aligned to your voting identity (*align*) reduces cooperation with Remainers. Consistent with our hypothesis, polarisation through strong affiliation to a group reduces prosocial behaviour in the dictator game.

Table 3.4: OLS regression Dictator Game by contribution

Dep. variable: Amount sent to receiver	Transfer Leaver		Transfer Remainer		Transfer UK	
	(1)	(2)	(3)	(4)	(5)	(6)
Leaver	0.212	0.046	-0.613***	0.033	-0.192	-0.206
	(0.188)	(0.257)	(0.200)	(0.266)	(0.154)	(0.231)
Beliefs (Leaver)	0.757***	0.748***	0.722***	0.731***	0.774***	0.784***
	(0.040)	(0.038)	(0.039)	(0.037)	(0.038)	(0.035)
Polarisation						
Vote proud		-0.014		0.010		-0.018
		(0.100)		(0.102)		(0.088)
Winner		-0.043		-0.727***		-0.222
		(0.219)		(0.232)		(0.206)
Align		0.312		-0.368**		-0.148
		(0.192)		(0.182)		(0.170)
Controls	NO	YES	NO	YES	NO	YES
Constant	0.951***	-2.748	1.665***	2.462**	1.191***	-1.783
	(0.130)	(2.207)	(0.223)	(1.085)	(0.190)	(2.505)
Observations	505	505	505	505	505	505

Note: Standard errors clustered at the subject level in parentheses.*** p<0.01,** p<0.05,* p<0.10.

3.4.3 Trust Game (TG)

Figure 3.11 shows the transfers and beliefs in the trust game. Beliefs are lower than transfers for each receiver, respectively. The average amount transferred to Leave receivers is 3.95 points. Remain voters, on average, received greater transfers at 4.73 points. UK receivers received the greatest points on average at 5.05 points. Figure 3.12 shows the two identities' divergence from the mean transfer, as indicated by the zero line. The figure shows that Leavers transfer more than the mean to fellow Leave voters and less than the mean to

Remain voters. Moreover, Remainers transfer greater than the mean to fellow Remainers and less than the mean to Leave voters. Both Leave and Remain participants are less trusting toward UK receivers.

To identify differences in trust by voters, we use Wilcoxon rank-sum test. When the receiver is a leaver or a Remainer, there is a statistically different level of trust between Leave and Remain participants, respectively (p =0.000 and p =0.000). However, there is no statistical difference in trust when the receiver is a UK citizen by Leavers or Remainers (p =0.987). Therefore, we find evidence of taste-based discrimination between Leavers and Remainers. However, the national identity is perceived as a non-rival group as there are no behavioural differences between Leavers and Remainers towards British receivers.

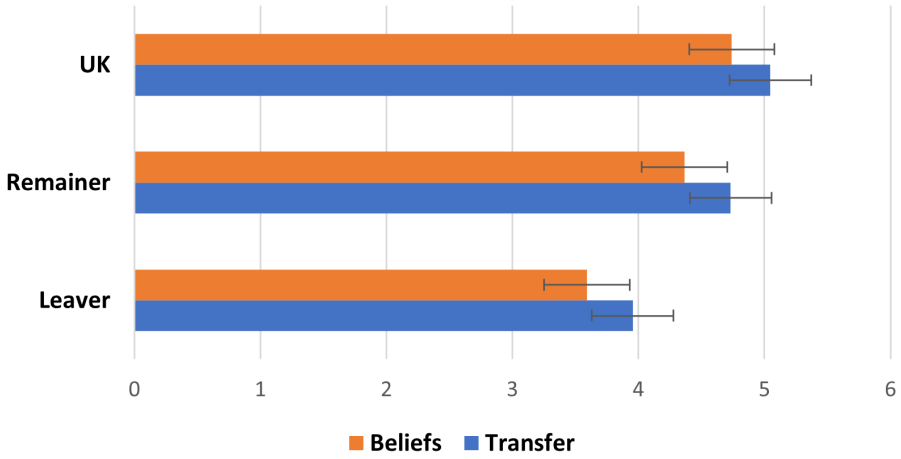


Figure 3.11: Trust Game (Mean contributions and beliefs)

Note: The figure shows the mean average transfers and beliefs in the Trust Game. The orange bars present beliefs, and the blue bars present the amount sent for each receiver (UK, Remainer and Leaver). Error bars indicate 95% confidence intervals.

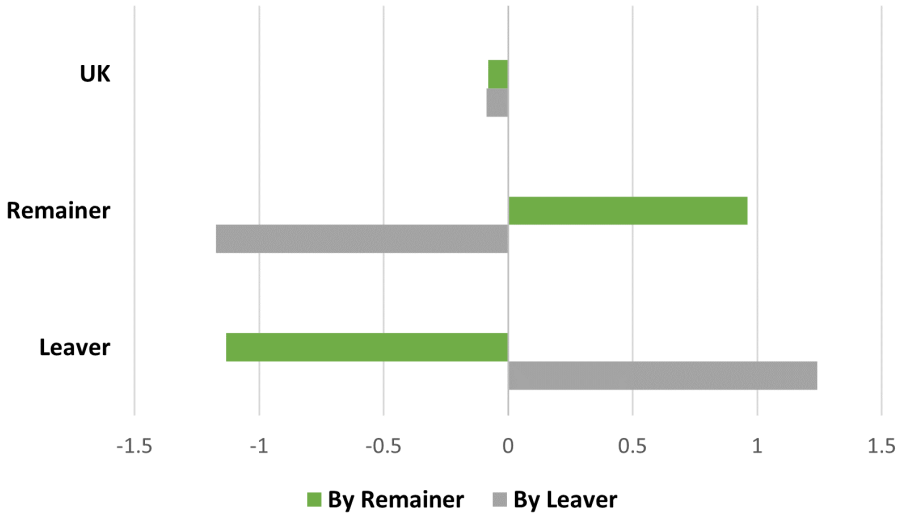


Figure 3.12: Trust Game (mean-centred)

Note: Zero line shows the mean benchmark. The orange bars show differences from mean transfers by Remain voters. The blue bars show transfers different from the mean by Leave voters.

We then used OLS regression, where the dependent variable is the amount sent to each receiver. We also control for age, income, sex, educational attainment, age, household size, ethnicity, race, marital status and risk-seeking, see Table 3.5. Leave trustors are less trusting towards Remainers. However, this is not robust to the inclusion of controls. The results show that beliefs about the receiver significantly impact trust and cooperation. Consistent with our hypothesis, *being proud* of the voting decision reduces coordination when the receiver is a Remainer. Feeling like a *winner* increases trust with Leavers and decreases trust with Remainers.

Table 3.5: OLS regression Trust Game by contribution

Dep. variable: Amount sent to receiver	Transfer Leaver		Transfer Remainer		Transfer UK	
	(1)	(2)	(3)	(4)	(5)	(6)
Leavers	0.129	-0.182	-0.459**	-0.020	-0.014	0.001
	(0.208)	(0.292)	(0.225)	(0.321)	(0.149)	(0.232)
Beliefs	0.844***	0.841***	0.736***	0.730***	0.878***	0.880***
	(0.033)	(0.032)	(0.038)	(0.038)	(0.024)	(0.025)
Polarisation						
Vote proud		0.019		-0.124		-0.006
		(0.101)		(0.129)		(0.094)
Winner		0.447*		-0.713***		-0.212
		(0.230)		(0.255)		(0.198)
Align		0.049		-0.054		0.159
		(0.189)		(0.213)		(0.171)
Controls	NO	YES	NO	YES	NO	YES
Constant	0.864***	0.749	1.736***	0.846	0.892***	1.063
	(0.130)	(1.536)	(0.254)	(3.030)	(0.155)	(0.959)
Observations	505	505	505	505	505	505

Note: Standard errors clustered at the subject level in parentheses.*** p<0.01,** p<0.05,* p<0.10.

3.4.4 Public Goods Game (PGG)

Figure 3.13 shows the transfers and beliefs in the public goods game. Beliefs are lower than transfers for Leaver, Remainer and UK receivers. The average amount transferred to Leave receivers is 3.83 points. Remain voters, on average, received greater transfers at 4.68 points. UK receivers earned the greatest points on average at 4.91 points. In Figure 3.14 zero indicates the mean contributions for all receivers. The figure shows that Leavers transfer more than the mean to fellow Leave voters and less for Remainers and UK

receivers. Conversely, Remainers transfer greater than the mean to fellow Remainers and UK receivers.

To identify differences in altruism and cooperation by voters, we use Wilcoxon rank-sum test. We observe a similar pattern as the other games. When the receiver is a leaver, there is a statistically different level of cooperation between Leave and Remain participants (p=0.000). Similarly, when the receiver is a Remainer, there are statistically different levels of cooperation between Leave and Remain participants (p=0.000). However, there is no statistical difference in trust when the receiver is a UK player by Leavers or Remainers (p=0.230).

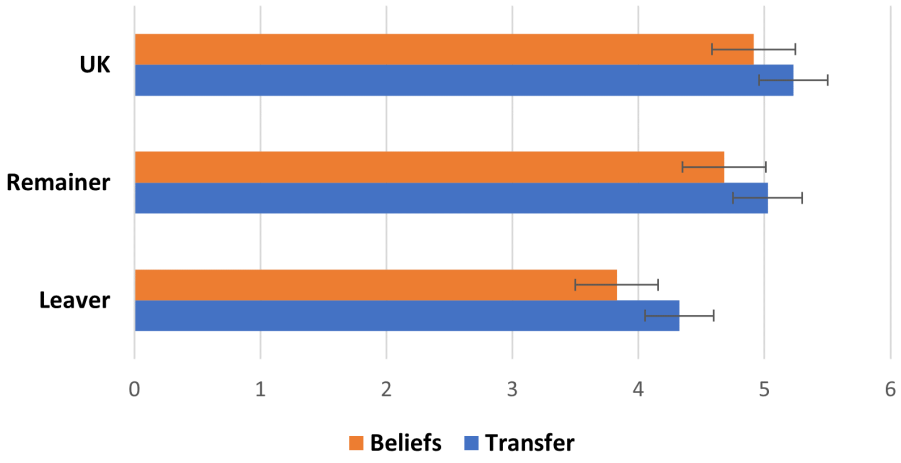


Figure 3.13: Public Goods Game (Mean contributions and beliefs)

Note: The figure shows the mean average transfers and beliefs in the Public Goods Game. The orange bars present beliefs, and the blue bars present the amount sent for each receiver (UK, Remainer and Leaver). Error bars indicate 95% confidence intervals.

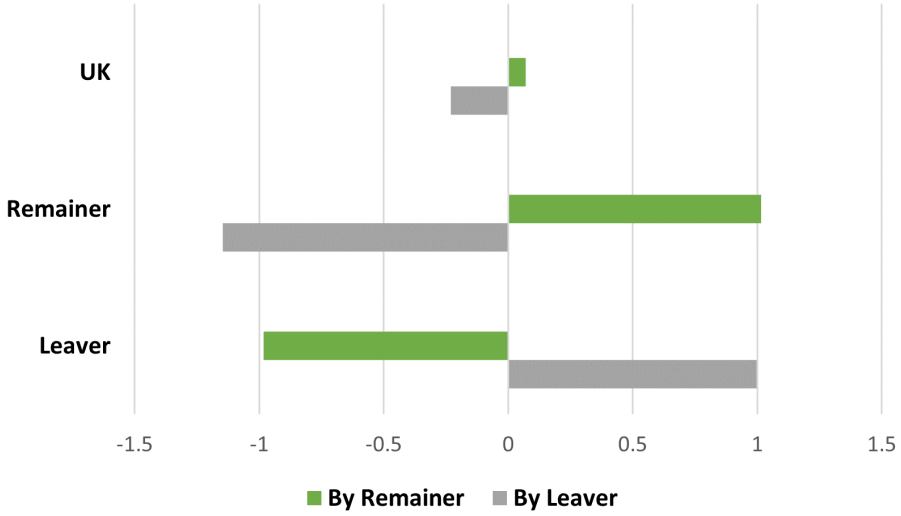


Figure 3.14: Public Goods Game (mean-centred)

Note: Zero line shows the mean benchmark. The orange bars show differences from mean transfers by Remain voters. The blue bars show transfers different from the mean by Leave voters.

Table 3.6 reports OLS regression in the public goods game. Beliefs regarding the respective player significantly influence cooperative behaviour. The greater belief in each player, the greater the amount sent, respectively. Consistent with our hypothesis, feeling like a *winner* reduces cooperation with Remainers. In addition, being highly aligned (*align*) with identity increases transfers to Leavers.

Table 3.6: OLS regression Public Goods Game by contribution

Dep. variable: Amount sent to receiver	Transfer Leaver		Transfer Remainer		Transfer UK	
	(1)	(2)	(3)	(4)	(5)	(6)
Leaver	0.129	-0.081	-0.626***	-0.131	-0.124	-0.094
	(0.197)	(0.273)	(0.208)	(0.276)	(0.141)	(0.231)
Beliefs (Leaver)	0.844***	0.838***	0.774***	0.768***	0.869***	0.870***
	(0.029)	(0.029)	(0.032)	(0.033)	(0.023)	(0.024)
Polarisation						
Vote proud		-0.136		-0.066		-0.054

Table 3.6 – continued from previous page

	Transfer Leaver		Transfer Remainer		Transfer UK	
	(1)	(2)	(3)	(4)	(5)	(6)

		(0.100)		(0.112)		(0.079)
Winner		0.083		-0.833***		-0.083
		(0.255)		(0.256)		(0.234)
Align		0.355*		-0.038		0.079
		(0.205)		(0.181)		(0.162)
Controls	NO	YES	NO	YES	NO	YES
Constant	1.032***	1.831	1.698***	3.745***	1.018***	2.417***
	(0.144)	(1.435)	(0.217)	(1.307)	(0.145)	(0.854)
Observations	505	505	505	505	505	505

Note: Standard errors clustered at the subject level in parentheses.*** p<0.01,** p<0.05,* p<0.10.

3.4.5 Modified Public Goods Game (mPGG)

Figure 3.15 shows the transfers and beliefs in the modified public goods game. Beliefs are lower than actual transfers. The average amount transferred to Leave receivers is 3.98 points. Remain voters, on average, received greater transfers at 4.70 points. UK receivers received the greatest points on average at 5.14 points. In Figure 3.16 zero indicates the mean contributions for all receivers. The figure shows that Leavers transfer more than the mean to fellow Leave voters and less for Remainers and UK receivers. Conversely, Remainers transfer greater than the mean to fellow Remainers and UK receivers and less than the mean to Leave receivers.

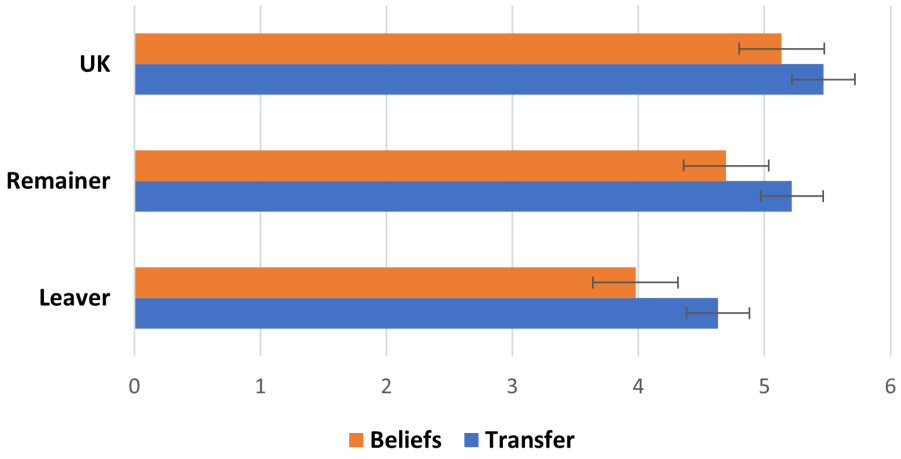


Figure 3.15: Modified Public Goods Game (Mean contributions and beliefs)

Note: The figure shows the mean average transfers and beliefs in the modified Public Goods Game. The orange bars present beliefs, and the blue bars present the amount sent for each receiver (UK, Remainer and Leaver). Error bars indicate 95% confidence intervals.

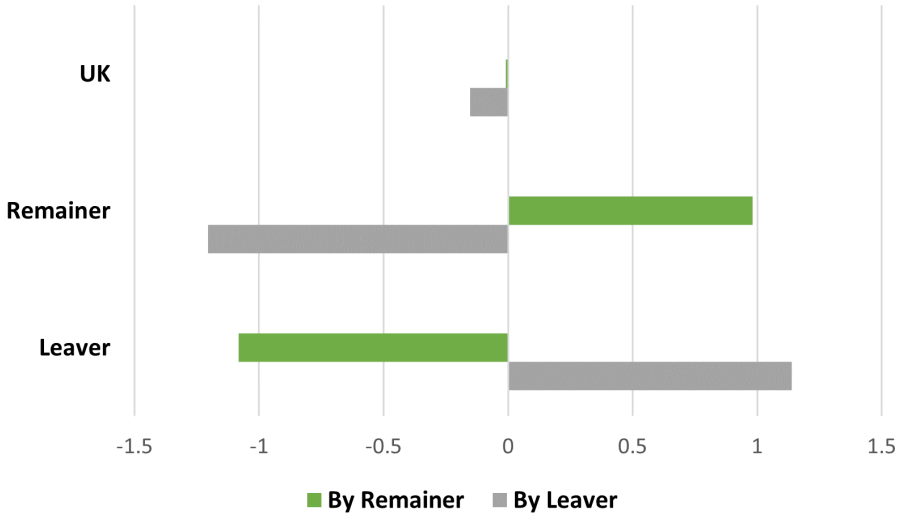


Figure 3.16: Modified Public Goods Game (mean-centred)

Note: Zero line shows the mean benchmark. The orange bars shows differences from mean transfers by Remain voters. The blue bars shows transfers different from the mean by Leave voters.

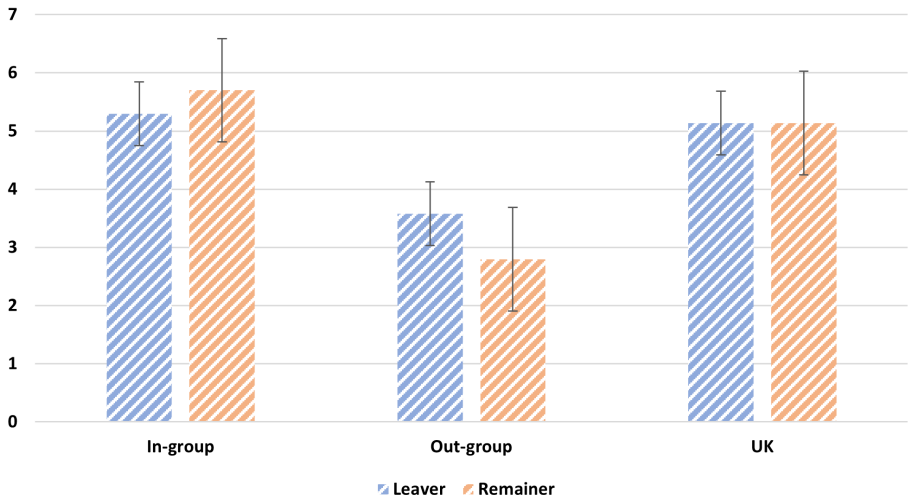


Figure 3.17: Modified Public Goods Game (mean-centred)
Note: Reported beliefs by Leavers (blue) and Remainers (orange).

Table 3.7 presents OLS regression for the modified public goods game controlling for age, income, sex, educational attainment, age, income, household size, ethnicity, race, marital status and risk-seeking. The dependent variable is the amount sent to receivers. The results show that beliefs about the receiver significantly impact coordination. Interestingly, feeling like a *winner* reduces the amount sent to British receivers [7].

Table 3.7: OLS regression (modified) Public Goods Game by contribution

Dep. variable: Amount sent to receiver	Transfer Leaver		Transfer Remainer		Transfer UK	
	(1)	(2)	(3)	(4)	(5)	(6)
Leaver	0.184	-0.013	-0.437**	-0.118	-0.140	0.083
	(0.201)	(0.260)	(0.178)	(0.281)	(0.156)	(0.219)
Beliefs	0.815***	0.811***	0.824***	0.821***	0.858***	0.860***
	(0.029)	(0.030)	(0.027)	(0.029)	(0.025)	(0.025)

Table 3.7 – continued from previous page

	Transfer Leaver		Transfer Remainer		Transfer UK	
	(1)	(2)	(3)	(4)	(5)	(6)

[7] We also estimated the same set of regressions using session fixed effects rather than individual effects. The results are unchanged. As a further robustness test, we repeat the set of regressions using a Tobit regression model to account for censoring, which yields similar results as OLS regressions, see Appendix B Table B1 - Table B3

Polarisation						
Vote proud		-0.071		-0.104		-0.018
		(0.117)		(0.129)		(0.096)
Winner		0.175		-0.268		-0.409**
		(0.223)		(0.250)		(0.203)
Align		-0.051		-0.253		-0.029
		(0.206)		(0.184)		(0.167)
Controls	NO	YES	NO	YES	NO	YES
Constant	1.303***	1.877	1.558***	3.629**	1.127***	2.952
	(0.159)	(1.628)	(0.205)	(1.411)	(0.181)	(2.124)
Observations	505	505	505	505	505	505

Note: Standard errors clustered at the subject level in parentheses.*** p<0.01,** p<0.05,* p<0.10.

The results of discriminatory behaviour show that identity matters. We find evidence of out-group discrimination. Interestingly, participants who voted to Leave are less cooperative with the in-group, although this finding is insignificant with the inclusion of controls. Notably, we observe a similar pattern in the Trust Game and Public Goods Game. We posit that this is explained by the negative coefficients of *vote proud* and *align*. Specifically, highly proud Leavers and those who align with their voting decisions are less cooperative with the in-group. Here, it is worth noting that aligning with one's voting decisions does not necessarily equate to aligning with the identity group. Furthermore, this pattern can, in part, be explained by the beliefs. For both Leavers and Remainers beliefs are greater for the in-group than the out-group. However, across all games, Leaver's beliefs of the in-group are lower than Remainer's beliefs of their in-group in absolute terms, suggesting lower expectations even with the in-group, see Figure 3.17.

In general, we find that the non-rival group, UK players, do not lead to in-group favouritism and out-group discrimination. Therefore, we observe discriminatory behaviour in opinion-based groups solely. Results from the self-reported polarisation measures reveal that feeling proud of one's vote, feeling like a winner and being highly aligned to one's identity reduces cooperation regardless of the receiver's identity. Overall, there was a widespread belief that

all players would contribute less than actual transfers. Although, across all economic games, there was the general belief that British players (or the UK) contribute more than people who did not vote, Remainers or Leavers. Across all games, there is a belief that Leavers will contribute less than Remainers. That said, we consistently find that beliefs significantly determine behaviour.

3.4.6 Social norms elicitation task

To analyse the effect of social norms on behaviour, we now report findings from the social norms elicitation task. We take a similar approach to Krupka & Weber (2013). To begin, we converted participants' responses into numerical scores. Participants who evaluate an action to be "very socially inappropriate" is valued at -1, "somewhat socially inappropriate" is valued at -1/3, "somewhat socially appropriate" is valued at 1/3 and "very socially appropriate" is valued at 1. Table 3.9 to Table 3.12 presents participants evaluations of social appropriateness. The first three rows show the total social norms. The following rows divide the sample by whether the participant is a Leaver or a Remain voter.

Table 3.8 reports the findings on the modal response of vandalising a Leavers home. The table shows that the modal response of this behaviour is that it is *very socially inappropriate*. We find no significant difference in evaluations across all majority and individual perceptions. Therefore, we find strong conformity between Leavers and Remainers concerning vandalising a Leaver's home. In Table 3.9, which shows participants modal evaluation of each action's social appropriateness. The table shows that the modal response is very socially inappropriate for all participants. We compare differences in social norms between participants who voted to leave and those who voted to remain using Wilcoxon rank-sum test. We find significant differences in individual perceptions regarding vandalising a Remainers garden (p =0.033) between Leavers and Remainers. Perceptions regarding the majority of Leavers are statistically different (p =0.000) between Leavers and Remainers. Similarly, we find evidence of differences between Leavers and Remainers regarding the majority of Remain voters (p =0.056).

In the social norms scenario regarding marriage between Leave and Remain voters. We find significant differences in individual perceptions (Wilcoxon rank-sum test p =0.000). Perceptions regarding the majority of Leave voters are statistically different between Leave and Remain voters (Wilcoxon rank-sum test p =0.000). We find weak evidence that there is a statistically significant difference in perceptions of most Remain voters between Leave

and Remain voters (Wilcoxon rank-sum test p =0.092).

We compare the differences in social norms regarding reporting bribery between Leave and Remain voters, see Table 3.11. We find a significant difference in individual perceptions (Wilcoxon p=0.005). Most Remain voters' perceptions are statistically different between Leave and Remain participants (p=0.029). Turning to punishing bribery, see Table 3.12. Although the modal response is the same across both identities, there is a significant difference in individual perceptions (Wilcoxon rank-sum test p =0.015). Most Remain voters' perceptions are statistically different between Leave and Remain participants (Wilcoxon rank-sum test p =0.088). To sum up, the majority of evaluations differ between Leavers and Remainers.

Table 3.8: Elicited social appropriateness of vandalising a Leavers home

	−−	−	+	++	N
Leave majority	87.92	7.33	3.96	0.79	505
Own perception	91.88	5.74	0.99	1.39	505
Remain majority	63.96	23.76	8.51	3.76	505
Participant is a Leave voter					
Leave majority	87.87	7.11	3.77	1.26	239
Own perception	93.31	3.77	0.42	2.51	239
Remain majority	66.53	17.57	9.62	6.28	239
Participant is a Remain voter					
Leave majority	87.97	7.52	4.14	0.38	266
Own perception	90.60	7.52	1.50	0.38	266
Remain majority	61.65	29.32	7.52	1.50	266

Note: Responses are "very socially inappropriate" (–), "somewhat socially inappropriate" (-), "somewhat socially appropriate" (+), "very socially appropriate" (++). Modal responses are shaded.

Table 3.9: Elicited social appropriateness of vandalising a Remainers home

	−−	−	+	++	N
Leave majority	57.03	26.14	13.07	3.76	505
Own perception	93.66	3.76	1.19	1.39	505
Remain majority	87.72	7.52	2.38	2.38	505
Participant is a Leave voter					
Leave majority	69.04	20.50	7.11	3.35	239
Own perception	91.21	5.44	0.84	2.51	239
Remain majority	84.94	8.37	2.51	4.18	239
Participant is a Remain voter					
Leave majority	46.24	31.20	18.42	4.14	266
Own perception	95.86	2.26	1.50	0.38	266
Remain majority	90.23	6.77	2.26	0.75	266

Note: Responses are "very socially inappropriate" (–), "somewhat socially inappropriate" (-), "somewhat socially appropriate" (+), "very socially appropriate" (++). Modal responses are shaded.

Table 3.10: Elicited social appropriateness of marrying someone that voted differently in the Referendum

	−−	−	+	++	N
Leave majority	4.64	12.14	30.18	53.04	505
Own perception	3.93	2.50	15.54	78.04	505
Remain majority	5.18	11.43	29.46	53.93	505
Participant is a Leave voter					
Leave majority	4.95	12.08	29.11	53.86	239
Own perception	4.16	2.77	15.84	77.23	239
Remain majority	5.54	11.09	28.32	55.05	239
Participant is a Remain voter					
Leave majority	5.64	18.80	36.09	39.47	266
Own perception	3.38	4.14	22.18	70.30	266
Remain majority	4.51	12.03	33.08	50.38	266

Note: Responses are "very socially inappropriate" (–), "somewhat socially inappropriate" (-), "somewhat socially appropriate" (+), "very socially appropriate" (++). Modal responses are shaded.

Table 3.11: Elicited social appropriateness of reporting bribery

	−−	−	+	++	N
Leave majority	14.06	10.30	21.19	54.46	505
Own perception	16.83	4.55	11.88	66.73	505
Remain majority	14.65	6.53	17.03	61.78	505
Participant is a Leave voter					
Leave majority	17.57	8.79	15.48	58.16	239
Own perception	20.50	5.86	12.97	60.67	239
Remain majority	17.99	7.11	17.57	57.32	239
Participant is a Remain voter					
Leave majority	10.90	11.65	26.32	51.13	266
Own perception	13.53	3.38	10.90	72.18	266
Remain majority	11.65	6.02	16.54	65.79	266

Note: Responses are "very socially inappropriate" (–), "somewhat socially inappropriate" (-), "somewhat socially appropriate" (+), "very socially appropriate" (++). Modal responses are shaded.

Table 3.12: Elicited social appropriateness of punishing bribery

	−−	−	+	++	N
Leave majority	5.74	8.91	35.84	49.50	505
Own perception	5.35	4.55	29.50	60.59	505
Remain majority	4.95	6.34	33.47	55.25	505
Participant is a Leave voter					
Leave majority	5.86	7.11	35.98	51.05	239
Own perception	6.69	5.44	32.64	55.23	239
Remain majority	5.86	7.11	35.56	51.46	239
Participant is a Remain voter					
Leave majority	5.64	10.53	35.71	48.12	266
Own perception	4.14	3.76	26.69	65.41	266
Remain majority	4.14	5.64	31.58	58.65	266

Note: Responses are "very socially inappropriate" (–), "somewhat socially inappropriate" (-), "somewhat socially appropriate" (+), "very socially appropriate" (++). Modal responses are shaded.

Table 3.13: OLS Regression on effects of social norms on mean contributions in games

	PD	DG	TG	mPGG	PGG
	(1)	(2)	(3)	(4)	(5)
Leaver	0.423	-0.258	-0.153	-0.104	0.281
	(0.554)	(0.420)	(0.586)	(0.391)	(0.405)
Family norm	-0.026	-0.050	0.155	-0.221	-0.048
	(0.263)	(0.217)	(0.173)	(0.207)	(0.215)
Family norm x Leaver	-0.375	0.384	-0.167	-0.036	-0.020
	(0.374)	(0.284)	(0.241)	(0.285)	(0.297)
Vandalising Leaver	-0.041	0.099	0.435	-0.406	-0.662
	(0.664)	(0.455)	(0.384)	(0.392)	(0.538)
Vandalising Leaver x Leaver	0.584	0.495	0.203	0.302	0.989
	(0.794)	(0.665)	(0.648)	(0.502)	(0.694)
Vandalising Remainer	0.660	-0.487	-0.440	0.188	0.712
	(0.582)	(0.434)	(0.397)	(0.375)	(0.485)
Vandalising Remainer x Leaver	-0.699	-0.197	-0.005	-0.595	-0.510
	(0.691)	(0.655)	(0.535)	(0.487)	(0.628)
Reporting Bribery	0.050	-0.036	-0.028	0.310*	-0.113
	(0.152)	(0.156)	(0.147)	(0.163)	(0.155)
Reporting Bribery x Leaver	-0.390*	-0.058	0.150	-0.130	0.322*
	(0.209)	(0.194)	(0.184)	(0.207)	(0.193)
Punishing bribery	0.086	-0.098	0.030	-0.066	0.041
	(0.205)	(0.216)	(0.176)	(0.165)	(0.250)
Punishing bribery x Leaver	-0.193	0.285	0.326	0.118	0.067
	(0.284)	(0.301)	(0.233)	(0.222)	(0.294)
Controls	YES	YES	YES	YES	YES
Constant	5.775***	1.736*	0.450	1.418	0.838
	(1.300)	(0.996)	(1.081)	(0.958)	(1.019)
Observations	504	505	505	505	505

Note: The dependent variable is the pooled mean contributions for all receivers (leave, remain and the UK) in each game. We add demographic controls and beliefs. Robust standard errors clustered by session in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

To identify if there is any divergence between personal and group norms, we run a simple t-test. In the vandalism scenario on Leaver, we find that Remainers' views of their in-group perceptions and their own norms differ when evaluating vandalism on Leaver's gardens ($p = 0.000$). We observe a similar pattern for Leavers' personal and in-group norms when evaluating vandalising a Remainers garden ($p = 0.000$). In both instances, we observe a wider spread of evaluations for the in-group than personal norms. In sum, we find evidence of pluralistic ignorance where the individual's private preference differs from the group (Smerdon et al. 2020, Miller & Prentice 2016). See Appendix B for tests of pluralistic ignorance for each norm-based treatment.

Table 3.13 reports the OLS regression of mean contributions regressed on social norm perceptions. We use the mean evaluations of each scenario. Overall, we find limited impact of norm perceptions on mean transfers and contributions across all games. Therefore, social norms do not significantly dictate cooperative behaviour in terms of mean contributions. We replicate these findings using disaggregated norm evaluations by each evaluation, although an individuals own perceptions regarding the family scenario significantly impact cooperativeness, see Appendix B Table B10.

Overall, the results from the social norms elicitation tasks reveal that over several vignettes, the modal social appropriateness ratings are similar across choice environments - own evaluations, majority of Leavers, and the majority of Remainers. There is general agreement on the social appropriateness of actions. This is consistent with social identity theory (Akerlof & Kranton 2000, 2005). That said, we find evidence of pluralistic ignorance, and the results show a significant difference between individual perceptions and in-group norms. However, we fail to find a significant influence of social norms on behaviours.

3.4.7 Heal and Divide intervention

Turning now to the intervention, we examine the degree to which social norm-based interventions impact behaviours. We also identify if the social norm-based intervention influences the evaluations of the social norms scenarios. We compare our analysis to the *Control* treatment, where no information is provided that acts as a baseline against measuring susceptibility to social influence when informed of the prevailing social norms.

In Figure 3.18, we observe variations in the level of coordination across the different treatments in the prisoner's dilemma game. We observe less coordination in the *Heal* and *Divide*

treatments in comparison to the *Control*, however, this is not statistically significant. We compare the difference in coordination between the social norms treatments in the prisoner's dilemma and the voting decisions of participants using the Wilcoxon rank-sum. When the participant is a leave voter, there is a statistically significant difference in coordination between the *Control* and *Divide* treatment when the receiver is a Leaver (p =0.020). There is no significant difference in coordination when the receiver is either a Remainer or a UK player. Non-parametric tests reveal that when the participant is a Remainer, there is no statistically significant difference in coordination between the *Control* and treatments across all receivers.

We further evaluate how the social norms treatments impact coordination using Probit regression models, see Table 3.14, Panel A. The *Heal* treatment reduces coordination with British receivers. The *Divide* treatment reduces coordination with Leavers. In column (3), we include the interaction term between identity and the treatments. If the participant is a Leave voter, the *Divide* treatment decreases coordination with Leavers and British receivers. The findings suggest that emphasising differences between Leavers and Remainers through the *Divide* treatment reduces coordination inter-group but also decreases coordination with the non-rival British group.

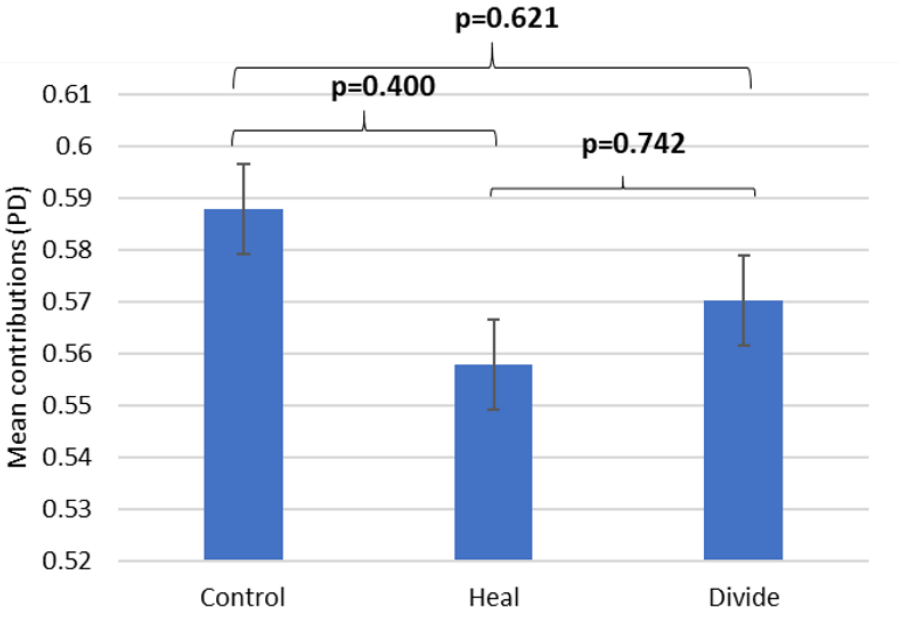


Figure 3.18: Prisoner's Dilemma by intervention

Note: The figure shows the fraction of coordination by the intervention treatment. Error bars indicate 95% confidence intervals.

In Figure 3.19, we observe that transfers are lower in the *Heal* and *Divide* treatments compared to the control. No statistical difference is found between mean contributions in the *Heal* or *Divide* treatments compared to the control (using non-parametric tests). Using Wilcoxon rank-sum test, we also compare cooperation by Leavers and Remainers following the treatments. The tests reveal no effect of the treatments when the participant is either a Leave voter or a Remain voter across all receivers. Using an OLS regression model, we compare the difference in contributions and transfers between the treatments in the dictator game with the inclusion of beliefs and controls. We find no evidence of treatment effects following the social norm-nudges in the dictator game. See Table 3.14, Panel B.

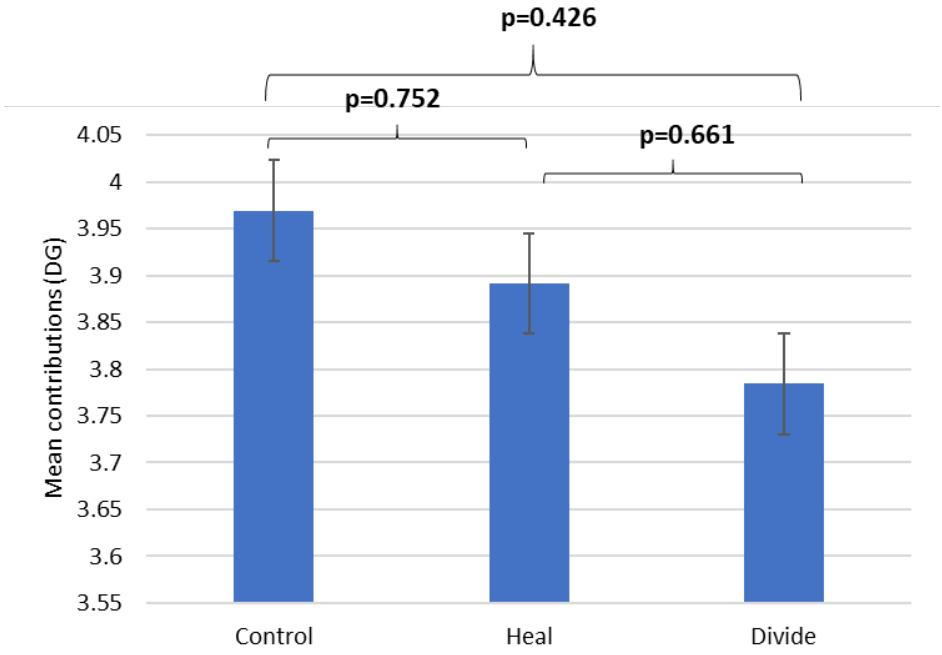


Figure 3.19: Dictator Game by intervention

Note: The figure shows the mean contribution in the Dictator Game by the intervention treatment. Error bars indicate 95% confidence intervals.

Figure 3.20, presents the amounts sent in the trust game in each treatment. Compared to the earlier games, we observe that the *Heal* treatment yields greater levels of trust than the *Control*. However, no statistical difference is found between mean contributions in the *Heal* or *Divide* treatments compared to the *Control*. Comparing trust behaviours between Leavers and Remainers using non-parametric tests, we find no statistically significant difference in cooperation in the trust game when participants are exposed to either the *Heal* or *Divide* treatment by either voting decision. Table 3.14, Panel C, reports OLS regression of treatment effects. We fail to find a significant effect of the social norm-based nudges on trust behaviours.

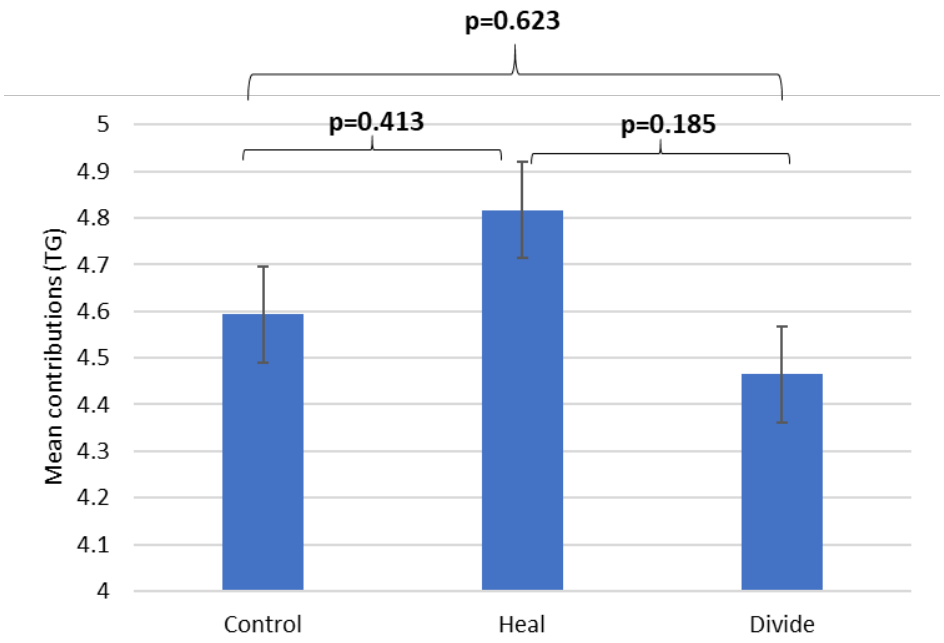


Figure 3.20: Trust Game by intervention

Note: The figure shows the mean contribution in the Trust Game by the intervention treatment. Error bars indicate 95% confidence intervals.

In the standard and modified public goods game, we observe that contributions in the *Heal* treatment are greater than in *Control* group. While contributions in the *Divide* treatment are less than the Control, see Figure 3.21 and Figure 3.22. In both instances, we find no statistical difference using t-tests. However, we are interested in identifying if there are any differences in behaviour following the treatments depending on the political voting decision made by the individual.

In the public goods game, following the *Heal* treatment relative to the *Control*, we find that if the participant is a Leave voter, there is a statistically significant difference in cooperation when the receiver is also a Leaver and Remainer (Wilcoxon rank-sum test $p = 0.054$, Wilcoxon rank-sum test p =0.035, respectively). Similarly, following the *Heal* treatment, when the participant is a leave voter, there is a statistically significant difference in cooperation when the receiver is the UK player (Wilcoxon rank-sum test p =0.0237). When the participant is a Remain voter, we find no statistically significant difference in cooperation between either treatment. We further test the treatments using an OLS regression model in Table 3.14, Panel D. In column (3), we include the interaction term of identity and treatments. The *Heal* and *Divide* treatments have negative impacts for Remainer receivers, when *Leaver equals zero*, therefore fore Remain participants. Moreover, if the participant voted to

Leave, the *Heal* and *Divide* lead to increased cooperation with Remainers. In this case, both treatments worked to facilitate increased cooperation from Leavers to Remainers. Therefore, the treatments backfire in the public goods game and cause in-group hostility between Remainers. Conversely, both treatments bridge the gap between Leavers and Remainers.

In the modified public goods game, the differences in behaviour by the voting identities are compared. When the participant is a Leave voter, following the *Heal* treatment, we find a statistically different level of cooperation across all receivers - Leavers, Remainers, and UK-citizens (Wilcoxon rank-sum test p =0.007, p = 0.017, p =0.001, respectively). No treatment effect is observed when the participant is a Remain voter. Table 3.14, Panel E reports OLS regression analysis of treatment effects. The *Heal* treatment increases cooperation for British receivers. If the sender is a Leaver, the *Heal* treatment increases cooperation for Remainers. Therefore, increases inter-group cooperation.

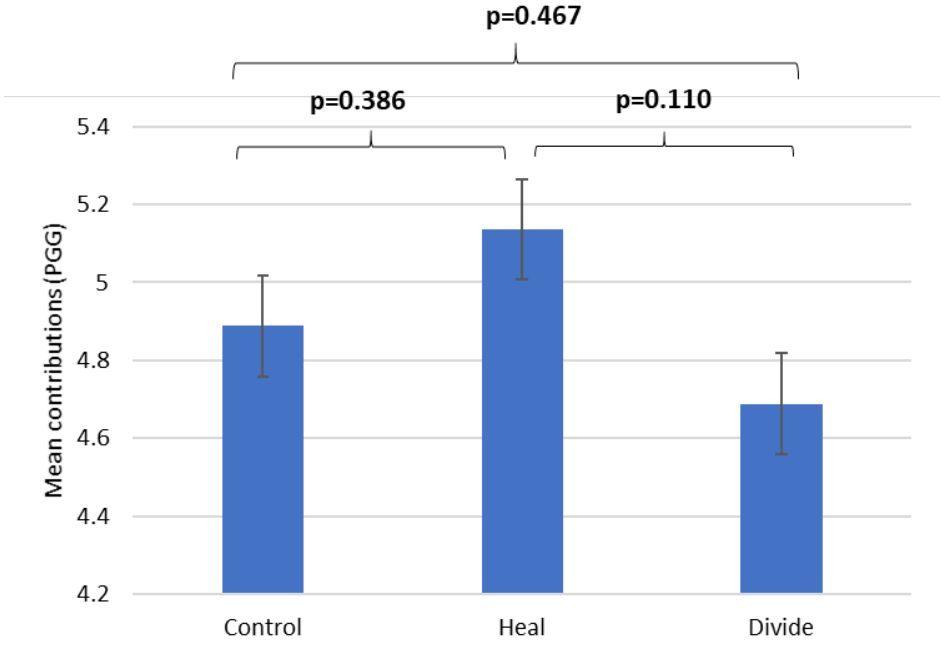


Figure 3.21: Public Goods Game by intervention

Note: The figure shows the mean contribution in the Public Good Game by the intervention treatment. Error bars indicate 95% confidence intervals.

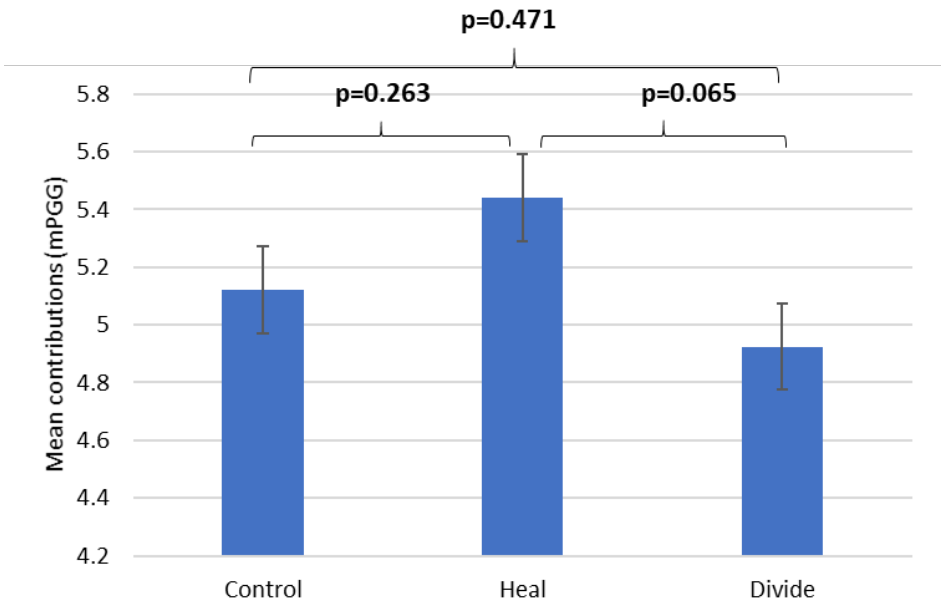


Figure 3.22: Modified Public Goods Game by intervention

Note: The figure shows the mean contribution in the Modified Public Goods Game by the intervention treatment. Error bars indicate 95% confidence intervals.

Table 3.14: Intervention effects on social norms

Dep. variable: Amount sent to receiver	Leaver			Remainer			British		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: Prisoner's Dilemma									
Heal	-0.122	-0.203	-0.135	0.083	0.106	0.130	-0.198	-0.277*	-0.093
	(0.131)	(0.149)	(0.202)	(0.131)	(0.151)	(0.218)	(0.134)	(0.156)	(0.205)
Divide	-0.228*	-0.350**	-0.053	0.105	0.219	0.254	-0.010	-0.089	0.191
	(0.130)	(0.147)	(0.194)	(0.130)	(0.147)	(0.217)	(0.134)	(0.152)	(0.199)
Leaver		0.305	0.589**		-0.478**	0.513**		0.214	0.547**
		(0.190)	(0.259)		(0.190)	(0.255)		(0.197)	(0.271)
Heal x Leaver			-0.198			-0.045			-0.425
			(0.308)			(0.307)			(0.319)
Divide x Leaver			-0.669**			-0.067			-0.629**
			(0.296)			(0.296)			(0.311)
Controls	NO	YES	YES	NO	YES	YES	NO	YES	YES
Constant	-0.096	5.248***	5.531***	0.391***	-5.457***	-5.940***	0.420***	1.018	1.463
	(0.108)	(0.670)	(0.705)	(0.110)	(0.678)	(0.708)	(0.112)	(0.914)	(0.954)
Observations	505	504	504	505	504	504	505	504	504
Panel B: Dictator Game									
Heal	-0.001	-0.050	-0.181	-0.128	-0.054	-0.251	-0.105	-0.206	-0.284
	(0.306)	(0.204)	(0.307)	(0.304)	(0.200)	(0.300)	(0.285)	(0.195)	(0.282)
Divide	-0.209	-0.114	-0.051	-0.215	0.082	0.094	-0.133	0.070	0.116

Continued on next page

Table 3.14 – continued from previous page

Dep. variable: Amount sent to receiver	Leaver			Remainer			British		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
	(0.294)	(0.200)	(0.265)	(0.294)	(0.194)	(0.231)	(0.267)	(0.182)	(0.236)
Leaver		0.046	0.006		0.036	-0.087		-0.195	-0.213
		(0.258)	(0.341)		(0.267)	(0.318)		(0.232)	(0.278)
Heal x Leaver			0.267			0.407			0.159
			(0.410)			(0.406)			(0.387)
Divide x Leaver			-0.127			-0.015			-0.094
			(0.399)			(0.390)			(0.371)
Controls	NO	YES	YES	NO	YES	YES	NO	YES	YES
Constant	2.433***	-2.806	-2.820	5.098***	2.515**	2.430**	4.320***	-1.725	-1.947
	(0.233)	(2.188)	(2.238)	(0.247)	(1.037)	(1.075)	(0.229)	(2.433)	(2.455)
Observations	505	505	505	505	505	505	505	505	505
Panel C: Trust Game									
Heal	0.081	-0.201	-0.155	0.280	0.191	0.246	0.310	0.146	0.397
	(0.347)	(0.218)	(0.310)	(0.340)	(0.243)	(0.333)	(0.316)	(0.193)	(0.245)
Divide	-0.110	-0.318	-0.398	-0.195	0.085	0.104	-0.080	0.013	0.098
	(0.327)	(0.200)	(0.261)	(0.317)	(0.200)	(0.252)	(0.299)	(0.170)	(0.205)
Leaver		-0.180	-0.212		-0.027	0.024		-0.005	0.221
		(0.293)	(0.380)		(0.321)	(0.392)		(0.233)	(0.295)
Heal x Leaver			-0.089			-0.115			-0.530
			(0.443)			(0.491)			(0.395)

Continued on next page

Table 3.14 – continued from previous page

Dep. variable: Amount sent to receiver	Leaver			Remainer			British		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Divide x Leaver			0.170			-0.044			-0.199
			(0.411)			(0.406)			(0.350)
Controls	NO	YES	YES	NO	YES	YES	NO	YES	YES
Constant	2.845***	0.595	0.405	5.670***	0.873	0.888	4.932***	1.056	1.229
	(0.264)	(1.513)	(1.538)	(0.258)	(3.042)	(3.061)	(0.252)	(0.936)	(0.986)
Observations	505	505	505	505	505	505	505	505	505
Panel D: Public Goods Game									
Heal	0.314	-0.103	-0.169	0.136	-0.040	-0.724***	0.291	0.120	-0.120
	(0.361)	(0.222)	(0.300)	(0.345)	(0.228)	(0.279)	(0.316)	(0.189)	(0.254)
Divide	-0.371	-0.440*	-0.347	-0.016	0.177	-0.496**	-0.216	-0.054	-0.201
	(0.336)	(0.230)	(0.310)	(0.335)	(0.210)	(0.237)	(0.305)	(0.173)	(0.230)
Leaver		-0.083	-0.054		-0.124	-1.068***		-0.101	-0.367
		(0.273)	(0.411)		(0.278)	(0.395)		(0.230)	(0.329)
Heal x Leaver			0.131			1.484***			0.513
			(0.449)			(0.471)			(0.393)
Divide x Leaver			-0.196			1.491***			0.330
			(0.463)			(0.429)			(0.351)
Controls	NO	YES	YES	NO	YES	YES	NO	YES	YES
Constant	3.455***	1.604	1.529	6.005***	3.844***	2.985**	5.334***	2.380***	2.068**

Continued on next page

Table 3.14 – continued from previous page

Dep. variable: Amount sent to receiver	Leaver			Remainer			British		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
	(0.290)	(1.394)	(1.449)	(0.258)	(1.274)	(1.251)	(0.257)	(0.813)	(0.851)
Observations	505	505	505	505	505	505	505	505	505
Panel E: Modified Public Goods Game									
Heal	0.372	0.005	-0.308	0.117	0.012	-0.451	0.466	0.395**	0.279
	(0.354)	(0.237)	(0.362)	(0.339)	(0.201)	(0.291)	(0.321)	(0.200)	(0.296)
Divide	-0.425	-0.281	-0.483	-0.185	0.171	0.064	0.018	0.303	0.252
	(0.335)	(0.232)	(0.355)	(0.331)	(0.196)	(0.246)	(0.303)	(0.195)	(0.272)
Leaver		-0.015	-0.361		-0.113	-0.509		0.071	-0.042
		(0.260)	(0.400)		(0.283)	(0.348)		(0.219)	(0.325)
Heal x Leaver			0.674			0.973**			0.245
			(0.463)			(0.406)			(0.412)
Divide x Leaver			0.453			0.251			0.114
			(0.461)			(0.398)			(0.398)
Controls	NO	YES	YES	NO	YES	YES	NO	YES	YES
Constant	3.632***	1.726	1.436	6.234***	3.719***	3.277**	5.360***	3.082	3.065
	(0.286)	(1.588)	(1.533)	(0.254)	(1.388)	(1.438)	(0.260)	(2.055)	(2.044)
Observations	505	505	505	505	505	505	505	505	505

Notes: Model (1) uses OLS regression estimation. Model (2) uses Ordered Probit regression estimation method. Robust standard errors clustered by session in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

In sum, our results provide evidence that social norm-based interventions influence behaviours. In group settings, such as the public goods game, we observe that the social norm-based nudge backfires and leads to less cooperativeness with in-group members between Remainers. However, for Leavers, the *Heal* and *Divide* treatments increase cooperation with Remainers. Therefore both treatments bridge the gap between diverging groups. However, it appears at the cost of in-group prosocial behaviour for Remainers. In addition, the *Divide* treatment decreases in-group coordination between Leavers and reduces coordination with British receivers in the prisoner's dilemma [8].

Turning now to the impact of the intervention of shifting evaluations of social norms, see Table 3.15. In column (1), we run an OLS regression, where the base category is the *control* group of the social norms intervention. Given the ordinal nature of the modal evaluations, we use the ordered Probit regression model in column (2). The dependent variable measures the appropriateness of actions. We expect that the *Heal* and *Divide* treatments would qualitatively decrease and increase social appropriateness ratings of the vandalism scenarios, respectively. In particular, we expect the vandalism scenarios to be consistent with this because of the similarities with the vignettes and the intervention treatments. We include the interaction of identity and treatments. The *Divide* treatments when *Leaver* equals zero, and therefore, for Remainers, lead to higher evaluations (social appropriateness) of vandalism in the three vandalism scenarios. However, the *Divide* treatment decreases modal evaluations (greater social inappropriateness) in the scenario regarding reporting bribery. However, if the participant is a leave voter, the *Divide* treatment leads to higher evaluations of reporting bribery. In other words, being informed that others view vandalism as socially appropriate increases the appropriateness of reporting bribery for Leave voters, but decreases the appropriateness of reporting bribery for Remain voters.

[8] We also estimated the same regressions but disaggregated by identity. The treatment effects remain relatively unchanged, see Appendix B, Table B5 - Table B9.

Table 3.15: Intervention treatment effects on behaviour

	Family		vandl		Vandr		vandno		Bribe 1		Bribe 2	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Leaver	0.047	0.287	0.111	0.158	0.044	-0.090	0.063	0.123	-0.244**	-0.292	-0.187**	-0.355*
	(0.085)	(0.209)	(0.070)	(0.222)	(0.071)	(0.223)	(0.062)	(0.238)	(0.123)	(0.202)	(0.089)	(0.190)
Heal	-0.012	0.080	0.002	-0.098	-0.011	-0.122	-0.003	-0.136	0.013	0.063	-0.100	-0.185
	(0.075)	(0.161)	(0.042)	(0.171)	(0.047)	(0.166)	(0.043)	(0.203)	(0.089)	(0.152)	(0.070)	(0.153)
Divide	-0.041	-0.085	0.090**	0.329**	0.085**	0.310**	0.081**	0.353**	-0.196**	-0.296**	-0.129*	-0.220
	(0.062)	(0.140)	(0.035)	(0.146)	(0.040)	(0.143)	(0.039)	(0.172)	(0.095)	(0.145)	(0.071)	(0.148)
Heal x Leaver	0.114	0.102	-0.080	-0.134	-0.032	0.036	-0.058	-0.103	0.034	-0.005	0.158	0.299
	(0.105)	(0.253)	(0.075)	(0.277)	(0.080)	(0.272)	(0.069)	(0.313)	(0.154)	(0.251)	(0.113)	(0.242)
Divide x Leaver	0.049	0.011	-0.133*	-0.323	-0.113	-0.219	-0.097	-0.183	0.321**	0.476*	0.183	0.314
	(0.101)	(0.245)	(0.068)	(0.243)	(0.071)	(0.246)	(0.064)	(0.264)	(0.153)	(0.244)	(0.114)	(0.237)
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	0.155		-0.573***		-0.553**		-0.682***		0.564		0.503	
	(0.320)		(0.221)		(0.265)		(0.219)		(0.472)		(0.354)	
N	505	505	505	505	505	505	505	505	505	505	505	505

Notes: Model (1) uses OLS regression, Model (2) uses Ordered Probit regression model. Robust standard errors clustered by session in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

3.5 Discussion and conclusion

Political polarisation and divisions can seep into other non-political behaviours in the UK, which in turn can hinder social cohesion and democratic practices. This paper investigates whether the referendum has led to the antipathy between Leavers and Remainers. Several years after the referendum, we find that individuals behave differently depending on if the person voted to Leave or Remain in the EU. Polarisation goes hand-in-hand with the growing populist rhetoric and movement (McCoy et al. 2018). Kuper (2022) reports that polarisation in the United Kingdom appeared to have peaked around 2016 to 2018. Whereas, in the United States, polarisation has lingered beyond Trump. However, we argue that polarisation following the referendum remains a salient issue.

We find evidence that individuals discriminate based on political identities, even in non-political contexts. In line with the conceptualisation of polarisation outlined in this paper, we find that alternative groups do not affect behaviour (Esteban & Ray 1994). Specifically, the neutral group or a British player has a limited effect on behaviour, and there is no difference in behaviour towards the British player between Leavers and Remainers (Bowles & Polania-Reyes 2012). The difference in behaviour between the out-group and the "non-rival" mixed-group supports findings by Abbink & Harris (2019), who similarly find less significant discriminatory behaviour when the player is in the mixed-group. Therefore, this provides a potential avenue for anti-discrimination policies to focus on national identity.

We use a modified version of the public goods game with two players to test the group effects that can cause polarisation (Cason & Mui 1997)—comparing the differences between the modified two-player public goods game and the Standard four-player public goods game. We find evidence of group effects that impacts discriminatory behaviour and the propensity for polarisation. We find greater shifts in behaviour following the norm-based interventions in the public goods game compared to the modified public goods game. Research by Cason & Mui (1997) and Bicchieri et al. (2022) find evidence that social proximity and group size are significant in enforcing norm compliance, discrimination and polarisation. Our findings support this lemma when looking at the behaviour in the standard versus modified public goods game.

Another explanation for discriminatory behaviour is stereotyping, as beliefs are found to significantly determine contributions across games. Consistent with the literature (see, for example, (Bigoni et al. 2019)), our analysis reveals that the beliefs about these different groups significantly impact cooperative behaviour. On average, reported beliefs are less than actual transfers; this suggests that perhaps the narrative surrounding the characteristics of Leavers and Remainers leading up to and following the Referendum (Hobolt et al. 2021, Virdee & McGeever 2018, Dalton 2019) has influenced opinions. In other words, the beliefs reveal that incomplete information leads to stereotyping or statistical discrimination, indicating group polarisation, as people have more negative perceptions or ideas about the opposing group.

Our self-reported measures of polarisation show that polarisation does negatively impact behaviour. We consistently observe that participants who highly regard themselves as 'winners' following the 2016 Referendum are less cooperative or altruistic. Our finding complement experimental research on political ideology and identities; our polarisation measure offer explanations for why others such as Thomsson & Vostroknutov (2017) and Anderson & Mellor (2009) fail to find significant results. The extent of the association or salience is necessary to understand inter-group behaviours of political and opinion-based identities.

Recent literature has shown that elections themselves may also shift evaluations of social norms (Apffelstaedt et al. 2022). In the social norm elicitation task, we find similar evaluations across scenarios by Leavers and Remainers. Therefore, given the relative proximity of the 2016 referendum and then the eventual leaving of the EU, this may also have influenced social norms evaluations. However, the social norms vignettes used in this study did not significantly affect overall cooperative behaviour across the experimental games. Our finding is consistent with Iyengar & Westwood (2015), who similarly observe discriminatory behaviour occurs along partisan lines without establishing strong social norms. The absence of influencing norms may be further exaggerated due to the relatively recent formation of our opinion-based identities.

Political polarisation has many negative consequences. For example, political polarisation damages social cohesion and public discourse and can put doubts on the validity of institutions (Sunstein et al. 2018, Dimant 2021, McCoy et al. 2018). When political campaigns lean into existing divides, this can often pull people further apart(Monroe et al. 2000). Conversely, in some instances, polarisation can increase participation in democratic practices through respectful debates amongst citizens (McCoy et al. 2018, Talisse 2022). However, in cases where polarisation hinders cooperativeness outside of the political arena and in day-

to-day interactions, polarisation can be harmful. Our research also endeavours to identify a remedy for political polarisation to support social capital (Nannicini et al. 2013). Therefore, the study also sought to investigate how (polarised) individuals can be further divided or how to bridge the gap between groups.

In our experiment, participants are randomly assigned to social norms intervention treatments. Following the social norm-based intervention, we find that the *Divide* treatment influences the evaluations of social norms by increasing the social appropriateness of vandalism. Our finding supports Albornoz et al. (2021) by providing evidence that vandalism (or hate crimes) are becoming more common or at least more socially accepted. We find inconsistent evidence regarding the effectiveness of social norm-based interventions in shifting behaviour towards more cooperative altruistic outcomes or more rational self-interested outcomes. Notably, both interventions increase inter-group interactions for the better. On the other hand, both interventions reduce in-group altruism between Remainers. An explanation for the mixed results following the social norm-based norm may be due to the frequent negative rhetoric and misinformation leading up to and after the Referendum (Iyengar & Westwood 2015). Therefore, information on Leaver's and Remainer's may be disregarded. Likewise, Su (2022) finds that an individual's political beliefs and perceptions are slow to update [9]. Nonetheless, the social norm-based nudge used in this paper could, for instance, be used in combination with other policies to increase social cohesion and capital alongside other policy motives (Barak-Corren 2021, Howlett et al. 2020).

In the previous paper, we looked at deep-rooted polarisation between two ethnolinguistic groups in Zimbabwe. There we found some evidence of in-group favouritism and out-group discrimination. Self-reported messages such as proudness and alignment also significantly impact behaviours in Zimbabwe. This is indicative of the polarisation and also the depth of the identities. Comparing the newly-formed opinion-based groups in the United Kingdom, we find evidence of in-group favouritism and out-group discrimination. Notably, we find comparatively weaker responses to the norm-nudge and the impact of social norms on cooperative behaviour in the UK compared to the experiments undertaken in Zimbabwe. In both experiments, discriminatory behaviour and in-group favouritism may be further displayed with different framing, for instance, games being framed as "community games" (Liberman et al. 2004, Dufwenberg et al. 2011). Contrary to Kuper (2022), this suggests that the divide post-Brexit is not wholly transient. It is interesting to note that we find evidence of pluralistic ignorance in both instances. The results from these studies reveal divergence

[9] See chapter 2 for general discussion of weakness of norm-based nudges.

between group norms and personal norms. As noted in chapter 2, pluralistic ignorance can be a mechanism that leads to more extreme behaviours. Which feeds into media discourse and negative press, which has been found to be one of the underlying causes that further divide groups, inter alia, that causes movement to the extremes (Smerdon et al. 2020, Miller & Prentice 2016, Bursztyn et al. 2020).

We acknowledge some limitations of the research. Political voting decisions have been explicitly identified in our experiment, unlike race which can be signalled visually. It is unclear how frequently the identity group of Leaver or Remainer is considered in decision-making. That said, research has shown that political identities do influence behaviours (Abbink & Harris 2012, Mill & Morgan 2021), however, further research is needed to understand the saliency of political identities and the complex interplay of aligning with the voting decision and with one's group. Future research could also consider the role of discourse and discussion in dampening polarisation (Sunstein 2019*a*, Hendriks et al. 2019).

Further research on the impacts of social norm-based nudges may prove fruitful, given the mixed results and debates around the use of nudges (Huitema et al. 2018, Lodge & Wegrich 2016, Ölander & Thøgersen 2014, Tyers 2019). However, Tyers (2019) highlights that some of the criticisms of nudges are mirrored in the broader political system, as there is currently a lack of trust and increased doubt in policymakers. Hence, policymakers have the challenge of ensuring that policies are trusted and consistent with the people's demands. For example, Sunstein (2019*b*) finds that there are differences along party lines in the United States regarding nudges and that, in general, informational campaigns received the greatest support across diverse countries (Sunstein et al. 2018), although the authors find that there is more approval of nudges in industrialised countries. Therefore, the solution to polarisation will need to be nuanced. Group polarisation influences democratic practices and is often used in political rhetoric, which can further polarised society - creating a potentially destructive cycle (Talisse 2022, Carlin & Love 2013).

In conclusion, we set out to investigate if there are social implications of the rising political polarisation in the United Kingdom following the 2016 EU referendum in terms of discriminatory behaviour. We find that Leavers and Remainers display inter-group discriminatory behaviour. Although this is a relatively new identity group that cannot be seen, the decision made in the referendum has created distinct groups that individuals relate to and, therefore, are willing to maximise their in-group benefits and maximise the difference between the in-group and out-group.

4

Would I lie for you?

Abstract

We experimentally investigate the extent to which people are willing to lie on behalf of someone else. We measure intrinsic (dis)honest behaviour using a variation of the Mind Game by Kajackaite & Gneezy (2017) to understand the propensity to lie for others when it also benefits the liar. We also test the role that social norms around the acceptance of dishonest behaviour moderate dishonest behaviour (Feldhaus & Mans 2014, Bucciol & Piovesan 2011, Mazar et al. 2008), using a variation of the social norms elicitation task (Krupka & Weber 2013). We conducted two experiments in the United Kingdom focusing on Leave and Remain voters. The second experiment also investigates under what conditions people are willing to lie for themselves and on behalf of others, using participants in Zimbabwe from the two major ethnic groups - the Shona and Ndebele. Firstly, the results are consistent with the literature that over ten rounds, participants do not maximise their payoffs in both contexts (Mazar et al. 2008, Khalmetski & Sliwka 2019, Erat & Gneezy 2012). Secondly, Leavers lie more than Remainers for themselves. In Zimbabwe, Ndebele participants lie more than Shona participants when the payoff goes to the participant only. Thirdly, we show that salience to a group identity influences the likelihood, emphasising the moral cost of dishonesty. Lastly, we find that the evaluation of social norms regarding reporting and punishing bribery significantly affects dishonest behaviour in both experiments. We contribute to the existing literature regarding dishonesty and shared identities providing an opportunity for corruption and bribery (Weisel & Shalvi 2015). This paper also contributes to the experimental literature on corruption and dishonest behaviour using a variation of the Mind Game.

Keywords: Dishonesty; Identity; Mind Game; Social norms

4.1 Introduction

Understanding people's willingness to lie is essential in economic research as dishonest behaviour can lead to the breakdown of contracts and an increased prevalence of crime and corruption (Cohn et al. 2019). There are growing concerns about dishonesty as state leaders and other influential persons' dishonest behaviour is exposed to varying degrees. For example, the reported wrongdoings of former President Donald Trumps leading to impeachment hearings (Congress of the United States of America 2019). Similarly, in the UK, Boris Johnson was exposed for lying about saving £350m a week on school dinners (Hattenstone 2020) and having several parties during the COVID-19 lockdowns (Cabinet Office 2022), and the continued accounts of corruption that are reported in Zimbabwe by Emmerson Mnangagwa (LeBas & Munemo 2019). Observing bribery and dishonesty can have negative economic consequences as citizens become more accepting of unethical behaviours. Therefore, in this paper, we experimentally test under what conditions people are willing to lie for themselves and on behalf of someone else, using participants from the United Kingdom and Zimbabwe.

The standard economic theory assumes people will be dishonest when the material incentive of lying outweighs the material incentives of being honest (Becker 1968). Previous experimental research has tested this theory using a form of a Deception Game or Cheating Game, such as sender-receiver games (Gneezy 2005), coin-flip tasks (Bucciol & Piovesan 2011), die-roll tasks (Fischbacher & Föllmi-Heusi 2013) and matrix-tasks(Mazar et al. 2008), amongst other forms. These experiments generally use private information and the opportunity for participants to behave dishonestly, which leads to a higher material payoff. The results of these experiments have observed dishonest behaviour; however, usually, participants are partially dishonest. Participants fail to maximise their material payoff. One of the leading explanations in the experimental literature for why individuals lie partially is that individuals have an aversion to lying; individuals will prefer to avoid being dishonest, even if being dishonest yields a positive payoff. Therefore, individuals balance the payoff from dishonest behaviour versus the disutility from lying (Mazar et al. 2008, Gneezy et al. 2018, Khalmetski & Sliwka 2019, Erat & Gneezy 2012). An alternative explanation for why partial lying is observed is due to social image concerns (Mazar et al. 2008, Fischbacher & Föllmi-Heusi 2013, Dellavigna et al. 2017, Grolleau et al. 2016, Ali & Bénabou 2016). People are honest because of self-image concerns and, therefore, the desire to be seen as doing the right thing.

As we are interested in investigating for whom someone is willing to lie, we also draw on group identity literature (Akerlof & Kranton 2000, Tajfel et al. 1971, Akerlof & Kranton

2005). We consider the role of affinity to a group in dishonest behaviours. When social links are involved in decision-making, agents are less likely to lie due to social ties (Feldhaus & Mans 2014). In contrast, other papers find the alternative to be true, that social relations increase cheating when it benefits someone with a shared social identity (Cadsby et al. 2016, Weisel & Shalvi 2015). This study expands on Erat & Gneezy (2012) research on Pareto White lies, which arise when the lie benefits them and someone else. Our paper is most closely related to research by Wiltermuth (2011) and Benistant & Villeval (2019). However, we deviate from this paper as the payoff from dishonest behaviour is not shared equally; this is because we aim to disentangle an individual's propensity to lie from themselves and their willingness to 'one up' for their in-group. Therefore, we contribute to the literature by examining group identity and its relationship with dishonest behaviour.

We also contribute to the literature on dishonest behaviour by conducting our experiments in developed and developing country settings. It is often argued that developing countries face a higher prevalence of corruption and that this trickles down to the rest of society (Di Zheng et al. 2020, Ryvkin et al. 2017). However, the issue of dishonesty from public officials is not limited to developing countries, only. Many significant Public officials in the West have been caught lying (Hattenstone 2020). In addition, Abeler et al. (2019) find that dishonest behaviour is similarly observed across countries. This paper uses the recently formed political identities of Leavers and Remainers in the UK (Hobolt 2016, Hobolt et al. 2021) [1]. In the second part of our experiment, we use pre-existing identity cleavages, the deep-rooted ethnic divisions in Zimbabwe, Shona and Ndebele [2].

Finally, our study also uses social norms elicitation tasks (Krupka & Weber 2013) in which participants are asked to evaluate socially appropriate behaviour in scenarios that involve dishonesty. This allows us to test if cheating is correlated with different norms surrounding the social appropriateness of honesty and dishonesty. Specifically, we conduct the following experiment: (i) *Norms* before undertaking the mind game, perceptions regarding dishonesty in a bribery scenario are elicited in a social norms elicitation task. In this scenario, the social appropriateness of reporting and punishing bribery is compared between the identity groups. (ii) *Mind Game treatments*, our experiment uses an adapted form of the mind game (Kajackaite & Gneezy 2017). Employing a between-subject variation in the games. This paper focuses solely on the Mind Game and a modified Mind Game to investigate deceptive

[1] See chapter 3 for further background into the UK context.

[2] Note that we do not set out to conduct a controlled cross-country comparison of dishonesty. However, we aim to identify the propensity and mechanisms surrounding dishonesty in both countries out of independent significance

behaviours. The between-subject design suffers from a potential lack of consistency, due to the inherent participant variability. Conversely, a strength of the between-subject research design is the ability to isolate the effects of specific conditions on behaviours (Charness & Gneezy 2012). Therefore, to minimise any noise, we maintain the same experimental environment of the Mind Game to clearly disentangle a participant's intrinsic willingness to lie for themselves and on behalf of someone else.

The results show that participants lie partially. In the experiments conducted with British participants, we find that Leavers report "Yes" more frequently than Remainers when the payoff benefits themselves only. We also find that salience towards the identity reduces dishonesty behaviour. Similarly, in the experiments conducted in Zimbabwe, we observed reduced dishonesty among in-group members. However, dishonest behaviour is based on national rather than ethnic identity. In both experiments, we observe that social norms regarding reporting and punishing bribery significantly influence dishonest behaviour.

The remainder of this paper is organised as follows. In section 4.2, we discuss the related literature and findings. In section 4.3, we identify the theoretical framework and hypotheses. Then in Section 4.4, we present our experimental design and procedure. Lastly, Section 4.5 discusses the results and concludes our findings.

4.2 Dishonesty and types of lying

There are multiple schools of thought regarding dishonesty and the moral implications of lying. A utilitarian approach does not always perceive a danger of lying; if it were to benefit overall social utility, then the benefit would outweigh the harm incurred by dishonest behaviour (Bentham 1781). Conversely, a Kantian approach argues that lying is always harmful. Lying harms others and reduces the self-respect of the individual. There are no instances in which dishonesty is appropriate; instead, we have a duty always to be honest (Kant 1949). On the other hand, Gneezy (2005) highlights the theologian Martin Luther who argued that there are instances where lying can be acceptable. There have been varying philosophical and ethical views on lying and its potentially harmful impacts. This study focuses on the environment in which individuals are willing to be dishonest.

Traditional economic theory posits that agents will be dishonest whenever the material payoff is positive (Becker 1968). In contrast to traditional theory, partial lying is observed in experiments (Fehr & Fischbacher 2002, Gneezy et al. 2018, Khalmetski & Sliwka 2019).

Participants believe that it is more honest to lie some of the time rather than all the time. For instance, individuals lie partially in the die-roll task, where the payoff increases the higher the number rolled. The unobservability of dishonest behaviour explains why some individuals report a four instead of five, as it appears more credible (Fischbacher & Föllmi-Heusi 2013, Erat & Gneezy 2012, Gächter & Schulz 2016, Gneezy et al. 2018). Another important consideration of dishonest behaviour is the environment individuals find themselves in. For instance, people may have concerns about the credibility of the lie.

Several taxonomies for lying have been identified by Erat & Gneezy (2012), spiteful black lies, selfish black lies, Pareto white lies and altruistic white lies. Here we focus on white lies; there are two main types of lying, altruistic white lies, which occur when another individual benefits from the lie at the expense of the one lying. In contrast, Pareto white lies benefit both the one lying and another. Erat & Gneezy (2012), using a deception game, found that participants do not lie when it comes to Pareto white lies. However, participants are willing to lie in the case of an altruistic white lie. In a similar vein, the literature on corruption has looked at how partnerships formed can foster lowered ethical behaviour, where participants are given the opportunity to lie together, in cases when being dishonest can benefit someone else, especially someone they identify with (Weisel & Shalvi 2015). In our paper, we focus on Pareto white lies, as this is more consistent with dishonest behaviour and corruption that benefits both parties [3].

4.2.1 Dishonesty games

Several lines of evidence suggest that people are willing to be partially dishonest in several conditions: if there is a high payoff association with being dishonest (Boles et al. 2000), if the positive payoff from lying is shared (Wiltermuth 2011) if there is a chance of measurement error (Toma et al. 2008) if others have been seen to be lying (Gino et al. 2013), and the likelihood that the reported outcome could be true (Sutter 2019).

Previous investigations into dishonest behaviour have varied in methodological structure. Overall, the literature has primarily focused on three forms of games when studying dishonesty - Deception Games, Cheating Games and Real Effort Games. Deception games are multilateral games that involve strategic decision-making. An example of a deception game is the Sender-Receiver Game (Gneezy 2005, Dreber & Johannesson 2008, Erat & Gneezy

[3]Specifically, we use a form of Pareto white lie. The participant lying receives a fixed positive benefit from lying; however, this is not proportionate to the benefit received by other individuals.

2012, Kajackaite & Gneezy 2017, Sutter 2019), the sender can send information regarding two options, with one being a false message and the other a truthful message. The receiver decides which option to take; the receiver will not know the actual payoff for each option. The decision made by the receiver will determine the overall payoff for both participants. The choice of deception is given to the sender as they choose between sending the truthful message or sending the false message; however, the wrong message yields a higher payoff. Whereas the truthful message, if the receiver trusts and follows up on it, will yield a lower payoff for the sender.

Alternatively, Cheating Games are non-strategic and thus do not involve beliefs regarding another player. In these games, a participant is given private information that only they will know and are asked to report this to the experimenter (the experimenter will never know the private information revealed to the participant) (Mazar et al. 2008, Fischbacher & Föllmi-Heusi 2013, Kajackaite & Gneezy 2017). Examples of cheating games include the die-roll and coin-flip tasks (see, for example, Arbel et al. (2014), Jacobsen et al. (2018), these game forms share many fundamental characteristics. In the die-roll tasks, participants can lie about the number that was rolled; for instance, in cases where the payoff increases as the number rolled increases, participants can lie by reporting a number slightly greater than the actual number observed. In the coin-flip tasks, they have to deliberately report the opposite of the observed coin-face to lie if the observed coin-face does not yield a positive payoff. Hence, the games differ because the die-roll allows for partial lying, as participants can report a four rather than a five to be seen as not being too dishonest (Gerlach et al. 2019, Bucciol & Piovesan 2011, Abeler et al. 2014).

In real effort games, the game incorporates the abilities of the participants. For example, in the matrix game used by Mazar et al. (2008), participants are tasked with solving several matrices. Participants are rewarded for correctly solving the matrices; therefore, they can be tempted to inflate the number of correctly solved matrices to gain a larger payoff. However, this game form suffers from a major issue as participants may truthfully misreport answers as they believe they have solved the matrices correctly (see Gerlach et al. (2019) for a review).

These game forms generally share the fundamental characteristic that participants have more information than the person who determines their payoffs. Participants are tempted into being dishonest as false information will yield a greater payoff[4]. The mind game used in

[4]The game forms are all positive-sum games where no one will lose out from dishonesty. Some studies have adapted these games and have used a constant sum game, whereby there is an externality from deception leaving an identifiable victim, see for example Erat & Gneezy (2012), Gneezy (2005). This paper primarily focuses on games with no identifiable victim resulting from dishonest behaviour.

this study can be seen as a combination of the two paradigms, as for each decision made, they have to decide to deliberately report "Yes" when it could be a "No". Moreover, at the beginning of the game, participants are made aware they will play over 10 rounds; therefore, over the 10 rounds, they may use a similar strategy outlined in the die-roll, for example, reporting "Yes" 6 times rather than 7 or more times. Unlike the matrix form, the modified Mind Game used in this study does not rely on the participant's skills. Therefore, eliminating some measurement errors that can occur in matrix games.

Previous research has manipulated different factors within these aforementioned experimental paradigms to observe dishonesty and understand if the degree of lying is greater in different environments. Several studies have investigated the effect of increased material stakes on dishonesty (Abeler et al. 2019, Fischbacher & Föllmi-Heusi 2013, Kajackaite & Gneezy 2017, Hugh-Jones 2016). The consequences of increasing the material payoffs depend on the experiment's game form. In deception games, the option of deceit is somewhat explicit in the game design (Hertwig & Ortmann 2008b). In cheating games, lying is more implicit; therefore, people may be more hesitant to lie (Jiang 2013, Grolleau et al. 2016). Alternatively, they may think they will get away with lying without the experimenter's awareness. Increasing the material payoff in cheating games may only make participants more concerned that the experimenter may observe their decisions and fear that any lying will be laid bare (Kajackaite & Gneezy 2017). In general, increasing the reward has had mixed effects on dishonesty in the experimental literature. This may be due to higher rewards increasing feelings of guilt (Mazar et al. 2008, Erat & Gneezy 2012), or individuals may perceive that higher rewards have a higher risk of being exposed (Kajackaite 2018). Alternatively, when the reward size induces more lying, it may be due to the higher reward leading to greater temptation (Gneezy 2005, Kajackaite & Gneezy 2017). In addition, higher stakes in the cheating game may backfire and cause less dishonesty because of the implicit nature of lying, whereas the alternative may be true for deception games (Gerlach et al. 2019).

Furthermore, field experiments have been utilised to explore deception in more natural settings (Bucciol & Piovesan 2011, Jacobsen et al. 2018). Cohn et al. (2019) investigate civic honesty in their cross-cultural experiment. In this field experiment, research assistants leave a wallet with an employee at a local counter. They waited to see if the email address was contacted to return the wallet; the amount of money in the wallet varied across treatments, thus altering the incentives. The authors apply this scenario in the field across 355 cities. Cohn et al. (2019) found that in most countries' citizens return the wallet when the wallet has money (behaved honestly), rather than a wallet without any money. They then increase

the stakes by having a treatment in which more money is in the wallet, although this only occurred in the United Kingdom, the United States and Poland. An explanation for why agents may be more honest in the field is due to fear of legal repercussions (Houser et al. 2016, Solaz et al. 2019), although this is not found to be significant in the study. Contrary to classic economic theory, the incentive did not increase dishonesty but rather increased civic honesty. Participants were more likely to return the wallet if the wallet contained a key (the key should be meaningless to participants). The key acted as a mechanism that highlighted altruistic concerns.

Dishonesty has generally been observed more in lab experiments than in field experiments. This may be because natural field experiments of this type are harder to control for exogenous factors. Given the nature of dishonesty and subsequently corruption and bribery, few individuals are willing to confess to willingly lying to benefit themselves and someone they share social ties with. In our study, we complement the findings from Cohn et al. (2019) by conducting experiments in different countries. Rather than in the field, we chose a controlled environment using a lab in the field experiment. Here, we also do not change the reward size for the individual but change the incentive scheme by incorporating the option for participants to lie on behalf of someone else. Experimental studies have shown that social ties lead to in-group favouritism and out-group discrimination across various contexts (Akerlof & Kranton 2005, Ostrom 1998, Tajfel et al. 1971, Krupka & Weber 2013, Rege, Mari; Telle 2004). Therefore, understanding when individuals are willing to be dishonest for and with is of economic importance.

Using a strategic real effort experiment, Benistant & Villeval (2019) explores whether the other player's identity in a competitive setting will impact the temptation to lie. Given the nature of the game paradigm used, shared group identities may alter unethical competitive behaviour (Cardenas et al. 2010, Abbink & Harris 2012, Benistant & Villeval 2019). The authors find that group affiliation is weaker than competitive drive (Benistant & Villeval 2019, Guala et al. 2013). Similarly, Di Zheng et al. (2020) found that bribery crowds out social affiliations between agents. However, our study proposes that in a non-competitive game paradigm, individuals may be tempted to commit Pareto white lies with in-group members rather than out-group members because of in-group favouritism, regardless of any strategic considerations. We focus on Pareto white lies because we can identify people's willingness to lie for others without confounding their willingness to lie for their material payoff.

4.2.2 Identity, normative cues and social norms

The economic and psychological literature has theoretically and empirically tested the factors influencing dishonesty. These factors have mainly focused on institutional trickle-down effects, identities, social and normative cues, inter alia (Di Zheng et al. 2020). People's views and perceptions regarding dishonesty are shaped and moulded by those around them, their culture and their proximity to dishonest or corrupt behaviour. Therefore, institutions where dishonesty is prevalent can influence the acceptance of dishonest behaviours (Bidner & Francois 2011, Birkelund et al. 2022, Gächter & Schulz 2016). Relatedly, a person's identity can also affect how they behave, whether this identity is their religion, gender, or career (Dreber & Johannesson 2008, Gerlach et al. 2019, Kocher et al. 2018, Shalvi & Leiser 2013). Non-monetary factors such as social norms can also be an essential determinant of lying, bribery, and corruption. These prevailing social ties can lead to dishonesty in favour of the in-group.

Dishonesty observed by someone who represents a specific social group can trickle down to the whole group, as individuals conform to the norm (Gächter & Schulz 2016). Additionally, emphasising distinct identities has been found to influence dishonest behaviour. Furthermore, reciprocity and social preferences, especially in group settings, can alter the propensity for dishonesty based on different groups (Feldhaus & Mans 2014, Bernhard et al. 2006, Eckel & Grossman 2004, Kocher et al. 2018). For instance, emphasising an individual's criminal history and status as a criminal leads to partial dishonest behaviour (Cohn et al. 2019). By focusing on their criminal identity, individuals will behave consistently with the perceptions of that identity. Solaz et al. (2019) found that individuals were willing to support and even forsake material payoffs for corrupt candidates if they had a shared identity. Wiltermuth (2011) found that individuals are more likely to cheat when the positive benefit of dishonesty is shared with another person. Broadly individuals use the shared benefit of dishonest behaviour to justify their dishonest behaviour (Erat & Gneezy 2012, Wiltermuth 2011, Feldhaus & Mans 2014). Therefore, social image concerns may be a principal determining factor in mitigating complete dishonest behaviour and thus can explain partial dishonest behaviour (Mazar et al. 2008, Fischbacher & Föllmi-Heusi 2013, Gächter & Schulz 2016, Grolleau et al. 2016). When focusing on political polarisation specifically, Mill & Morgan (2021) found evidence of dysfunctional and spiteful behaviour. We add to this literature by using a variation of the Mind Game to identify the extent that individuals may choose to be honest to ensure that the 'other group' loses out.

Given the extrinsic influence that can impact participants, normative cues can enforce the moral cost of dishonest behaviour. Normative cues work by either influencing the social norms around dishonest behaviour through informing participants about the behaviour of others or raising the salience of ethical behaviour. Using a combination of game forms, normative cues which provide information on the dishonest behaviour of other people can, in fact, increase the level of dishonesty (Bicchieri et al. 2021, Fosgaard et al. 2013, Abeler et al. 2019, Jacobsen et al. 2018). Conversely, emphasising ethical and moral norms across various game forms lowers dishonest behaviour (Mazar et al. 2008, Bucciol & Piovesan 2011, Shu et al. 2011).

Peer effects can change moral behaviour and preferences for dishonesty (Isler & Gächter 2022). A field experiment investigating dishonesty regarding voting turnout and voting behaviour has found that Chicago-based residents were willing to incur a lying cost to hide that they had not voted to avoid facing social punishment (Dellavigna et al. 2017). In this case, there is a cost of lying rather than a positive material payoff from lying, but the social image concern is great enough for agents to incur this cost. Thus, dishonesty has been observed even at a cost to the individual to preserve social image concerns.

Identities such as religious affiliations have been investigated concerning dishonest behaviour. Given the potentially greater emphasis on the moral implications of dishonest behaviour, scholars have compared religious participants with secular participants (Shalvi & Leiser 2013, Arbel et al. 2014, Isler et al. 2021). Although the results are mixed, Arbel et al. (2014) found that secular men and women cheat more than religious men and women when comparing Jewish participants with secular participants. Studies focusing on religious identity and lying suggest that religious individuals are less dishonest to maintain the image of appearing honest.

Gender differences in dishonest behaviour have also been found in the literature (see, for example, Dreber & Johannesson (2008), and Feldhaus & Mans (2014)). Moreover, there is some evidence that gender differences in cheating are partly due to individual worldviews that are hierarchical and individualistic (Birkelund et al. 2022). Hence, preferences for dishonesty can be explained by cultural factors; these can be broader cultural factors or closer to home. For instance, in a coin-flip game in the field experiment, Bucciol & Piovesan (2011) assesses dishonesty in children aged between 5 and 15 during an Italian summer camp. They found significant levels of cheating across ages and other characteristics like siblings. Notably, mentioning the need to be honest to children reduces dishonesty, but this is more effective in females than males. However, the strength of the request for honesty fell for

female participants with age but remained stable in male participants. Houser et al. (2016), also found that parents are dishonest when it benefits their children. However, parents are more honest when observed by their daughters than their sons.

Maintaining a moral and social image can enforce intrinsic motives. Dishonesty and greed can be seen as undesirable traits. Social image concerns have been identified in experiments where there is a possibility for scrutiny from an external person, whether the experimenter or another participant. To avoid appearing like a cheater, social image-focused people may exaggerate honest behaviour (Dai et al. 2018). Therefore, people do not maximise dishonest behaviour to appear morally correct (Mazar et al. 2008). Notably, theft aversion, an extension of the social image concern, is the specific self-image concern of not seeming to be a thief. This explanation was used by Cohn et al. (2019), as it was more important if the lost items were valuable to the person who found them rather than what would be more valuable to the person who lost the wallet. Feldhaus & Mans (2014) found that the moral cost of dishonest behaviour is far greater when individuals share an identity. To disentangle the motivations for own lying and preferences and willingness to lie for others, participants in the modified mind game treatments (Senders) received a fixed fee for dishonesty. Therefore, the receiver's payoff is decision-dependent on the sender. We can then compare dishonesty in the standard mind game, where the payoff is decision-dependent for the receiver.

Likewise, in experiments where participants could be observed, cheating is lower because of the social image concerns compared to non-observed games (Gneezy et al. 2018). Moreover, the timing and opportunity for lying effects dishonest behaviour. For instance, Jiang (2013) found that when the intent to cheat is made first, maintaining a moral social image is harder as individuals will intentionally delay reporting cheating. Hence, in this treatment group, where the intention to cheat is made at the beginning of the experiment, the authors found less cheating because of social image concerns. Therefore, we hypothesise that the social or moral cost of dishonesty will significantly impact behaviour when dishonesty occurs between individuals in the same identity group. To examine the influence of social norms around the 'right thing to do', we use a social norms elicitation task to elicit Krupka & Weber (2013) to identify an individual's norms around dishonest behaviour.

Therefore, will an awareness that someone with a shared identity will benefit from dishonesty lead to increased lying? Are people more likely to be honest when the payoff goes to an outgroup member as a means of spitefulness? Do people's normative perceptions of dishonest behaviour shape their behaviours?

To address these questions, this study varies the incentive scheme by creating a treatment where being dishonest benefits someone else. We correlate the propensity to be dishonest in the Mind Games with the perceived social norms of lying in different scenarios and reported opinions on faith in the government. All participants report their beliefs regarding dishonest behaviour in a social norms elicitation task. We use salient real-world social ties rather than artificial social ties. By inducing social identities, we can identify if differences in behaviour between groups are because of other characteristics outside of innate differences between the groups (Di Zheng et al. 2020). Thus, we are able to answer if bad institutions lead to a higher propensity for dishonesty. The originality of this study is that it delves into the environment people are willing to lie for others of their in-group and out-group, but also the strength of parochial altruism is identified by comparing the propensity to lie when participants are paired with a *mixed* group member, in this case, participants are aware that their partner is British or Zimbabwean which has a 50% chance of being either an in-group or out-group member.

4.3 Theoretical Framework and Hypotheses

Previous research on dishonest behaviour has posited that people will lie when the material incentive of lying outweighs the material incentives of being honest, regardless of the impact on others. Following Becker (1971), the literature has provided theoretical frameworks on the assumptions and preferences for honesty, whereby agents measure the moral benefits and costs of dishonesty (Kartik et al. 2014, Matsushima 2008, Ellingsen & Östling 2010). In contrast to classical models, research has found that agents do not always maximise their material payoffs by lying. A simple model of lying aversion by (Fischbacher & Föllmi-Heusi 2013) implies that agents get disutility from lying. However, this disutility can increase or decrease depending on the marginal material benefit of being dishonest. There is a trade-off between the material payoff gained from lying and the payoff from being honest. Participants with a falling disutility in the marginal benefit of lying will be dishonest when the material payoff of doing so is high enough compared to the payoff of being honest. Agents with rising disutility in the marginal benefit of lying will be dishonest until the point where the gain meets a specific threshold.

We use a simple utility function similar to Kajackaite & Gneezy (2017) to explain the decisions faced by an agent in the mind game and the modified mind game. Agents observe private information p and are tasked with reporting this information. The agent can either

be honest and report the true information p or be dishonest and thus report false information p'. The monetary payoff from being honest is m_p, and from being dishonest is $m_{p'}$. Hence, the monetary payoff of lying is $m_p - m_{p'}$. The intrinsic cost of lying is given by C_i. The fourth component is the social norms $\gamma_i j N(a)$ of action a.

$$m_p - m_{p'} - C_i - \gamma_{ij} N(a) > 0 \tag{4.1}$$

The function $\gamma(.)$ represents the social norms associated with dishonesty behaviour. The identity can be their self-identity i or the perceived identity placed on them by others j. $\gamma \geq 0$, as it increases, an individual will get higher utility from complying with the norm - in other words, doing the socially appropriate action. We assume that honest behaviour or being perceived as honest yields greater utility than dishonesty. Therefore, social categorisation and internationalisation of norms will influence dishonesty behaviours (Akerlof & Kranton 2000, 2005, Benistant & Villeval 2019). By identifying the social component of dishonest behaviour, people can register their decision as either negative or positive, consistent with the appropriate self-image consideration (Gino et al. 2013, Bicchieri et al. 2021). Furthermore, identity formation leads to in-groups and out-groups (Akerlof & Kranton 2000, 2005, Tajfel et al. 1971), which can influence the self-image concerns around dishonesty and dishonesty on behalf of others. Individuals behave strategically in response to incentives, which could involve drawing closer to a certain identity to justify dishonest behaviour (Posner 2005). Hence, the identity of the other individual is also of importance [5]. The benefit of dishonesty for others, especially an in-group member, can serve to justify cheating behaviours, and out-group discrimination may create a zero-sum mentality (McCoy et al. 2018). Therefore, individuals may be more likely to report the true information p to minimise the other players' payoff (Wiltermuth 2011). Conversely, conformity with the norm implies that if individuals believe the majority of the in-group disapproves of dishonest behaviour, then the moral cost of dishonest behaviour will be greater when interacting with an individual with a shared identity (Bicchieri et al. 2021, Feldhaus & Mans 2014).

Hypothesis 9 *Individuals will lie partially.*

Hypothesis 10 *In the Modified Mind Game, there will be a higher fraction of reported "Yes" outcomes when the payoff is distributed to someone with a shared identity.*

[5] See chapter 2 for further theoretical consideration of social identity theory.

Hypothesis 11 *Social norms will impact the propensity for dishonest behaviour.*

4.4 Experimental Methods

In the following section, we present the experimental design, the procedure and other measures of interest. We use a combination of lab-in-the-field and online experiments.

4.4.1 Social norms elicitation task

The social norm elicitation task occurred before the Mind Game to identify normative perceptions of amoral behaviour. In the social norm elicitation task (Krupka & Weber 2013), participants are described different scenarios and were asked to evaluate the social appropriateness of each action given the scenario. We use a within-subject design as all subjects do the social norm tasks before undertaking a variation of the Mind Game to reduce the consequence of subject-specific effects between beliefs regarding socially appropriate behaviour and actual behaviour in the mind game. Participants evaluated each scenario as either *very socially appropriate*, *somewhat socially appropriate*, *somewhat socially inappropriate* and *very socially inappropriate*. One of their answers was randomly selected, and they received a payoff depending on the accuracy of their guess. As participants were aware that their evaluation would be compared with another person's, if their assessment matched, they would earn 4 points [6]. The instructions can be found in the Appendix C.

The social appropriateness of lying and dishonesty were elicited in two scenarios. *Bribe Scenario 1* describes a scenario where someone reports a manager who is bribing an official public, and *Bribe Scenario 2* illustrates a scenario where a citizen can punish the manager who is bribing the public official. In each scenario, participants are presented with three cases. One case elicits what they deem socially appropriate, and the other identifies what they believe others in-group members would deem socially appropriate. A final case identifies what they believe others in the out-group would deem is socially appropriate [7]. In both experiments, we used the following social norms vignettes.

- *Bribe Scenario 1 - You see a manager bribing a public official to get a contract. How*

[6] In the UK, 4 points were converted to £2, and in the Second experiment, 4 points were converted to $2.
[7] see Appendix C for detailed instructions

appropriate is it to report the manager.

- *Bribe Scenario 2 - A citizen is made aware of this bribe. How appropriate is it to punish the manager.*

4.4.2 Standard Mind Game

To test dishonest behaviour, we use a variation of the Mind Game by Kajackaite & Gneezy (2017), Jiang (2013). Participants are asked to think of any number i between 1 and 10 and to remember this number, where $i \in \{1,2,3,4,5,6,7,8,9,10\}$. The Mind Game is used to measure non-strategic intrinsic levels of honesty, as the experimenter has no way of monitoring the simple decision made by the participant. Rather than rolling the dice as is the case in Kajackaite & Gneezy (2017), participants are provided with a randomly generated number that is shown on the screen. The participant is asked to report if the number they thought of is the same as the number randomly generated. If the participant reports `"Yes"`, then the participants receive a payoff of 1 point and zero; otherwise, this is repeated 10 times. Randomising the numbers shown and participants being able to choose the number privately provides participants with the opportunity to be dishonest. By repeating the game, lying is detected by the proportion of `"correct"` choices being statistically improbable.

4.4.3 Modified Mind Game

In this variation of the Mind Game, participants are asked to complete the same task as in the standard Mind Game, thinking of any number i between 1 and 10 and to remember this number, where $i \in \{1,2,3,4,5,6,7,8,9,10\}$. However, the payoff they earn is mirrored and given to someone. Participants are informed to whom the payoff will be sent. Participants will earn a fixed fee for taking part in this task [8]. Participants are randomly selected into groups where the payoff is sent to someone in their [voting or ethnic] group, someone of the other [voting or ethnic] group and someone who represents the national identity, British and Zimbabwean, respectively. The modified mind game remains a non-strategic game but incorporates social preferences as the sender is made aware of the other person's identity and who will receive the payoff; however, the receiver remains completely passive. We use a between-subject design to test our hypotheses. By varying the incentive settings, we can

[8]Participants receive a fixed fee to ensure they are incentivized. However, we did not match the payoff received by others for the player as the strategy would be very similar to the standard mind game

test the degree to which participants are willing to lie for themselves and others, either in their in-group, out-group or mixed. See Table 4.1 and Table 4.2 for payoff scheme of both experiments.

Table 4.1: Payoffs in the different treatments (UK)

Game variant	**Payoff Scheme**	**N**
Standard MG	Receive 1 point if the report is yes; 0 points otherwise	33
MG Leaver	player receives a fixed 4 points. If the report is yes, another Leaver player receives 1 point; otherwise, the other player receives 0 points.	47
MG Remainer	player receives a fixed 4 points. If the report is yes, another Remainer player receives 1 point; otherwise, the other player receives 0 points	52
MG British	Player receives a fixed 4 points. If the report is yes, another British participant receives 1 points; otherwise, the other player receives 0 points	41

Table 4.2: Payoffs in the different treatments (Zimbabwe)

Game variant	**Payoff Scheme**	**N**
Standard Mind Game	Receive 1 point if report yes; 0 points otherwise	52
Mind Game for Nd	player receives a fixed 4 points. If the report is yes, another Ndebele player receives 1 point; otherwise, the other player receives 0 points.	54
Mind Game for Sh	player receives a fixed 4 points. If the report is yes, another Shona player receives 1 point; otherwise, the other player receives 0 points	51
Mind Game for Zim	Player receives fixed 4 points. If the report is yes, another Zimbabwean player receives 1 point; otherwise, the other player receives 0 points.	52

4.4.4 Experimental Procedure

In the UK, the experiments were conducted online using the Prolific recruitment service. The experiment was programmed using the online survey software Qualtrics (Qualtrics 2020). The experiments conducted in Zimbabwe were over several sessions at Marondera Agricultural University, Harare Polytechnic, Nyadire Teaching College, Lupane State University, and the University of Zimbabwe. Our participants consisted of 26% Ndebele players and 74% Shona players. The majority of the subjects were students (79%). Students were randomly chosen from undergraduate and postgraduate courses. Participants joined the experiment

voluntarily by responding to an email sent to all students prior to the experimental sessions.

The experiment proceeded as follows. At the beginning of the session, participants received instructions. We asked participants to identify their voting decision in the 2016 EU referendum and their ethnic identity in the UK experiment and the Zimbabwean experiment, respectively. They were informed that they would receive a fixed participation fee and also an additional payoff which would differ for each participant. The additional payoff is determined by the number of times the participant guesses the correct number. If they guessed all 10 rounds correctly, they would receive 10 points. Before completing the mind game, participants undertook the social norm elicitation task. Participants are not given instant feedback regarding the social norm elicitation task. See Figure 4.1 for a summary of the experiment.

The players reporting of the number was truly private. In the instructions, it is made clear that the experimenter will not be able to verify if the number they chose in their head is, in fact, the same as the number shown on the screen. See Appendix C for instructions. The decisions were made in private, without the experimenter being able to detect if the participant was lying. This intends to reflect real-life situations in which corruption occurs because it can go unnoticed. Participants were paid for the mind game and social norms elicitation task after completing all tasks to avoid spillover effects. After completing the task, participants completed a post-experiment questionnaire that included gender, age, and income questions.

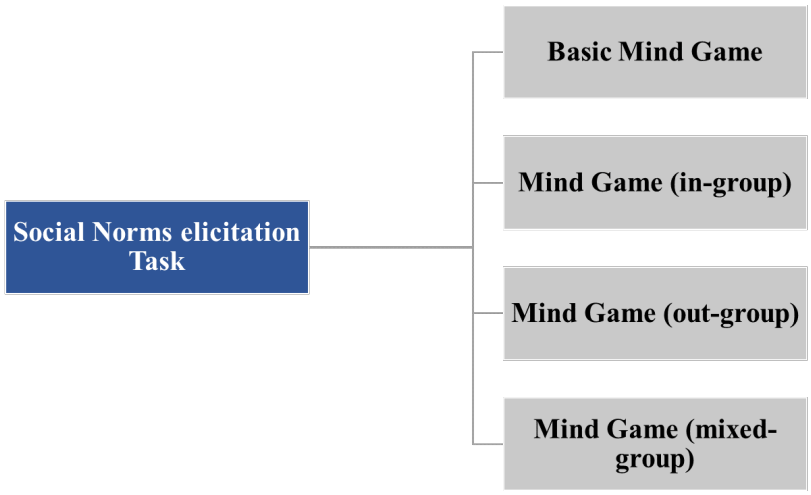


Figure 4.1: Summary of Experiments

4.5 Results 1 (UK)

Study 1 was conducted in the United Kingdom using the Prolific online recruitment platform. We began by investigating dishonest behaviour. In addition, we examined the level of dishonesty dependent on the identity of the player and the receiver. Lastly, we investigated the influence of social norms on dishonesty.

Table 4.3: Summary Statistics

	MG	**MG (Leave)**	**MG (Remaine)**	**MG (British)**
Identity				
Leaver (%)	45.45	36.17	48.08	48.78
Remainer (%)	54.55	63.83	51.92	51.22
Age (%)				
25-34	33.33	31.48	37.04	46.67
35-44	33.33	27.78	24.07	20.00
45-54	18.18	31.48	20.37	30.00
Female (%)	66.67	62.96	72.22	72.73
Single (%)	51.52	48.15	51.85	53.33
Household size	2.94	2.91	2.83	2.89
Risk seeking	4.36	4.20	3.85	4.04
Faith in gov	3.30	3.13	3.31	2.78
Risk with gov	4.73	4.67	4.24	4.09
Observations	33	47	52	41

Notes: Baseline summary statistics of participant characteristics in the experiment.

4.5.1 Identity and dishonest behaviour

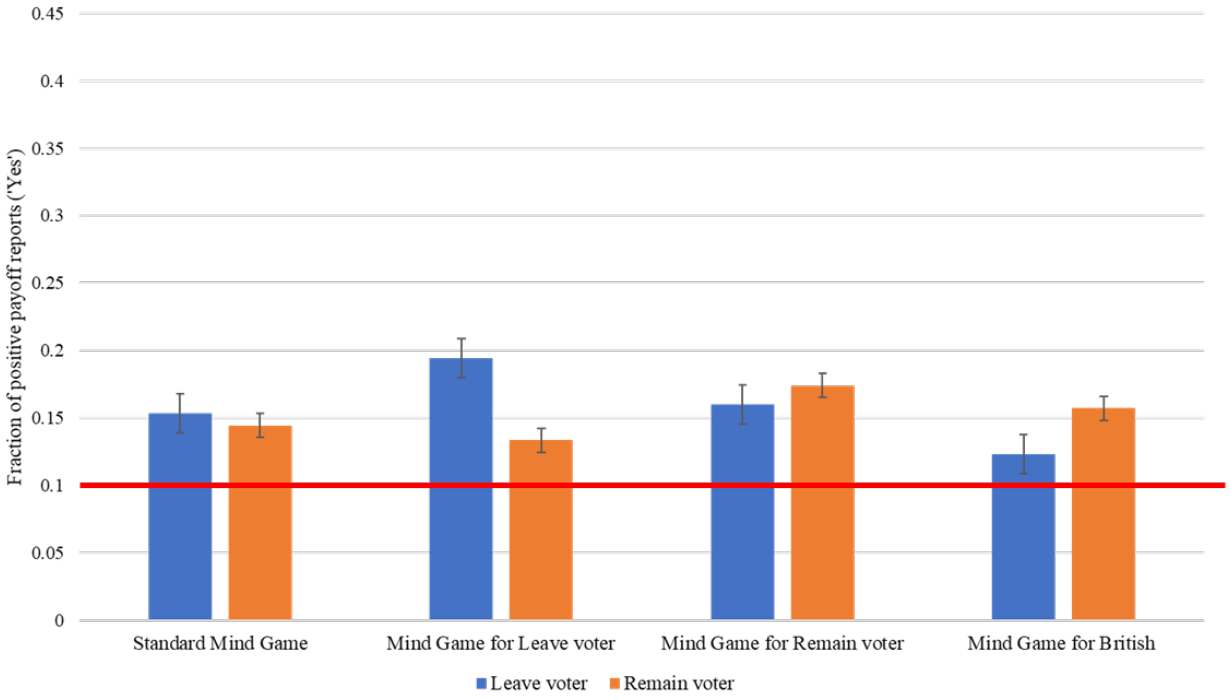


Figure 4.2: Lying the Mind Game in the UK

Note: Bars present the fraction of 'Yes' outcomes reported. The first set of columns presents the fraction of 'Yes' outcomes reported in the Standard Mind Game. The red line marks the expected 10% outcome. Error bars indicate 95% confidence intervals.

Figure 4.2 illustrates the fraction of participants who reported `"Yes"` - a positive payoff in the different treatments. The first two bars show the fraction of positive payoff in the Standard Mind Game over the 10 rounds, 15.33% of Leave voters report `"Yes"` and 14.44% of Remain voters report `"Yes"`. The second set of columns display the responses in the treatment where the payoff goes to Leave voters, 19.41% of leave voters report a `"Yes"` and 13.33% of Remain voters report a `"Yes"`. In the treatment where the payoff goes to Remain voters, 16.00% of Leave voters report a `"Yes"` and 17.41% of Remain voters report a `"Yes"` over the 10 rounds. In the treatment to the neutral group, a British participant, 12.50% of Leave voters report a `"Yes"` and 15.71% of Remain voters report a `"Yes"` over the 10 rounds.

Participants were tasked with guessing a number between 1 and 10 over 10 rounds. Therefore, the expected number of correct guesses (reporting `"Yes"`) is 10%. Thus, in this study, any value greater than 10% is considered to be due to dishonesty. Taking a binomial test shows

that overall in the standard Mind Game, there is a significant level of dishonesty above 10% ($p = 0.003$). When participants are Leavers or Remainers, there is a statistically significant level of reporting `"Yes"` ($p = 0.000$ and $p = 0.000$, respectively). Lastly, when the payoff goes to a British participant, there is a statistically significant fraction of `"Yes"` responses greater than 10% (bitest p =0.005). Although we find evidence of lying, our players display an aversion to lying, as no participant maximises their payoff.

Due to the binary nature of our dependent variable, we ran Probit regression models, where the base category is the control group Remainers, including controls for the demographic characteristics of participants. In column (1), we find that the likelihood of reporting `"Yes"` response is greater for Leavers than Remainers. In column (1m), the marginal effect of dummy *Leaver* shows that on average, Leavers reported `"Yes"` 16.5 percentage points more than Remain voters in the standard mind game. Therefore, in our Standard Mind Game, Leaveers are dishonest for themselves significantly more than Remain voters. We do not find evidence of greater Pareto white lies for Leavers and Remainers or British receivers.

Lying on behalf of others may be driven by other factors related to their voting decision. For instance, the variable *vote proud* indicates that individuals who are proud of their vote report fewer `"Yes"` outcomes in treatment 1, where participant solely benefits from dishonest behaviour and in treatments 2 and 4 where the payoff goes to a Leave voter and a British voter, respectively. Similarly, the dummy variable *winner* shows that those who feel like a winner following their vote lie less in the Standard Mind Game. Again, we observe a similar pattern with dummy variable *diffvote*, which equals one if the individual would vote the same given the opportunity again, and zero otherwise. Moreover, being highly aligned (*align*) to your voting identity leads to an aversion to lying for others when the other receiver is a Remain voter.

Interestingly, the variable *govtrust19* reports the amount that participants trust the government in their COVID-19 lockdown measures; higher values indicate greater trust. We observe that participants who are more trusting of government measures have a greater likelihood of reporting `"Yes"` outcomes when the other receiver is British. Whereas the variable *govtrust*, which indicates general trust in the government from 0 to 10, 10 being highly trusting of the government has no impact on dishonesty behaviour. We find that this does not alter dishonesty behaviours. As a robustness check, results are qualitatively similar when using random-fixed effects. See Appendix C Table C1 for robustness checks to account for decisions made over 10 rounds.

Table 4.4: Probit regression of dishonesty (UK)

Dep. variable: Reporting positive payoff	Standard		Modified (Leaver)		Modified (Remainer)		Modified (British)	
Model	(1)	(1m)	(2)	(2m)	(3)	(3m)	(4)	(4m)
Leaver(=1, Remainer=0)	0.778**	0.165**	0.079	0.017	0.260	0.061	0.044	0.009
	(0.346)	(0.076)	(0.264)	(0.058)	(0.233)	(0.053)	(0.210)	(0.042)
proud of vote	-0.200**	-0.040**	-0.152**	-0.032*	-0.095	-0.022	-0.073*	-0.014*
	(0.082)	(0.017)	(0.077)	(0.017)	(0.070)	(0.017)	(0.042)	(0.008)
winner	-1.166***	-0.236***	0.220	0.047	0.085	0.020	-0.165	-0.033
	(0.360)	(0.072)	(0.266)	(0.055)	(0.297)	(0.070)	(0.145)	(0.028)
diffvote	-0.648***	-0.131***	0.003	0.001	-0.376	-0.087	-0.030	-0.006
	(0.219)	(0.042)	(0.215)	(0.046)	(0.234)	(0.058)	(0.131)	(0.026)
align	0.557	0.113	-0.025	-0.005	-0.567**	-0.132**	-0.285	-0.056
	(0.340)	(0.069)	(0.194)	(0.041)	(0.229)	(0.053)	(0.180)	(0.035)
govtrust	0.004	0.001	-0.008	-0.002	0.017	0.004	-0.007	-0.001
	(0.051)	(0.010)	(0.040)	(0.009)	(0.043)	(0.010)	(0.028)	(0.006)
govtrust19	-0.092	-0.019	-0.027	-0.006	0.092	0.021	0.141***	0.028***
	(0.074)	(0.015)	(0.045)	(0.009)	(0.059)	(0.014)	(0.044)	(0.008)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Constant	3.280***		0.426		-1.030		-1.813***	
	(1.158)		(0.712)		(0.861)		(0.527)	
Observations	330	330	470	470	520	520	410	410

Notes: Column (1) presents Probit regression, and column (2) presents the marginal effects. Robust standard errors clustered at the subject level in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

4.5.2 Social preferences and dishonesty

To gain an understanding of the role social norms play in lying behaviours. All participants took part in the social norms elicitation task. We take a similar approach to Krupka & Weber (2013). To begin, we converted participants' responses into numerical scores. Participants who evaluate an action to be *very socially inappropriate* is valued at -1, *somewhat socially inappropriate* is valued at -1/3, *Somewhat socially appropriate* is valued at 1/3 and *very socially appropriate* is valued at 1. Table 4.5 and Table 4.6 present participants' evaluations of social appropriateness. We represent the findings in the three evaluated cases: their own perception, Leaver majority and Remainer majority.

Table 4.5: Bribe 1 scenario (UK)

	−−	−	+	++	N
Own perception	17.92	2.89	10.40	68.79	173
Leave majority	14.45	6.94	19.65	58.96	173
Remain majority	15.61	4.62	16.18	63.58	173
Participant is a Leave voter					
Own perception (L)	24.68	3.90	11.69	59.74	77
Leave majority (L)	23.38	6.49	11.69	58.44	77
Remain majority (L)	23.38	2.60	18.18	55.84	77
Participant is a Remain voter					
Own perception (R)	12.50	2.08	9.38	76.04	96
Leave majority (R)	7.29	7.29	26.04	59.38	96
Remain majority (R)	9.38	6.25	14.58	69.79	96

Notes: −− for ratings `"Very socially inappropriate"`, − for ratings `"Somewhat socially inappropriate"`, + for ratings `"Somewhat socially appropriate"`, and ++ for ratings `"Very socially appropriate"`. Modal responses are shaded. N is the number of observations.

Table 4.5 presents the reported social perceptions of *Reporting Bribery*. There are consistent modal evaluations across the four variants of norm evaluations. In this scenario, over 50% of participants evaluate reporting bribery as *very socially appropriate*. When the participant is a Leave voter (see rows 5-8), we observe similar average ratings when participants evaluate

Table 4.6: Bribe 2 scenario (UK)

	−−	−	+	++	N
Own perception	5.20	2.89	27.75	64.16	173
Leave majority	4.05	8.67	34.68	52.60	173
Remain majority	4.62	5.20	33.53	56.65	173
Participant is a Leave voter					
Own perception	7.79	6.49	32.47	53.25	77
Leave majority (L)	6.49	7.79	33.77	51.95	77
Remain majority	6.49	10.39	33.77	49.35	77
Participant is a Remain voter					
Own perception	3.13	-	23.96	72.92	96
Leave majority (R)	2.08	9.38	35.42	53.13	96
Remain majority	3.13	1.04	33.33	62.50	96

Notes: −− for ratings "Very socially inappropriate", − for ratings "Somewhat socially inappropriate", + for ratings "Somewhat socially appropriate", and ++ for ratings "Very socially appropriate". Modal responses are shaded. N is the number of observations.

the social norms based on the perceptions of the majority of leave voters, their perceptions and Remain voters. However, when the participant doing the evaluation is a Remain voter (see rows 9-12), the participant's own perceptions are greater than the evaluated majority perceptions.

We then compared the differences in social norms regarding reporting bribery between participants who are Leave and Remain voters. Using Wilcoxon rank-sum test, we find a significant difference in an individual's perceptions of reporting bribery between Leavers and Remainers (p=0.000). The perceptions of most Leave and Remain voters are evaluated statistically differently between Leave and Remain participants (p=0.001 and p=0.000, respectively). Therefore, these results suggest that the distribution of norm evaluations differs between Leave and Remain voters.

Table 4.6 presents the evaluations of the social appropriateness of *Punishing Bribery*. On average, punishing bribery is evaluated as being *very socially appropriate*. When the evaluator is a Leave voter, average evaluations are relatively similar when the evaluation is based on their perception and the majority of Leave voters and Remain voters. When the evaluator in the social norms task is a Remain voter, we observe different modal evaluations of punishing bribery. When evaluating their perceptions, 72.92% consider punishing bribery as *very*

socially appropriate. We find that Remainer's own perceptions are more pronounced than their in-groups and more so than Leavers. We find a significant difference in an individual's own perceptions between Leavers and Remainers (WC p=0.000). However, the perceptions of the majority of Leave voters are not statistically different between Leave and Remain participants (WC p=0.264). Most Remain voters' perceptions are evaluated as statistically different between Leave and Remain participants (WC p=0.000).

In Table 4.7, our regression models include the social norms evaluations. To identify how dishonesty is related to the social acceptance of bribery. Given the environment of the norms, norm-compliance to morally right action would lead to a lower likelihood of dishonest behaviour. Overall, we find that social norms around bribery influence dishonest behaviour. Consistent with our hypothesis, we find that the more socially appropriate reporting bribery is viewed by the majority of Leave voters, (*bribe1 (Leaver)*)the reduced likelihood of reported `"Yes"` outcomes in the Standard Mind Game and the modified Mind game where the payoff also goes to a Leaver. However, the more socially appropriate reporting bribery is viewed by the majority of Remain voters, *bribe1 (Remainer)*, reduces reported `"Yes"` outcomes in the Standard Mind Game, but increases the likelihood of dishonest behaviour when the other receiver is a Remainer. Furthermore, we observe some norm violations from the majority of Leavers (*bribe2 (Leaver)*) in the Standard Mind Game. Similarly, the perspective of the majority of Remainers (*bribe2 (Remainer)*) when the other receiver is a Remainer. Overall, we observe that social norm compliance is likely impacted by the identity of the recipient of dishonesty, which can lead to "moral leniency" (Shu et al. 2011, p.344) [9].

[9]See Appendix C, Table C3 for further robustness checks using Poisson regression models. This does not change the results. In addition, we replicate the set of regressions using the random-effects model to account for changes in behaviour over the 10 rounds. The results are unchanged. See Appendix C, Table C1

Table 4.7: Probit regression of dishonesty and social norms (UK)

Dep. variable: Reporting positive payoff	Standard		Modified (Leaver)		Modified (Remainer)		Modified (British)	
Model	(1)	(1m)	(2)	(2m)	(3)	(3m)	(4)	(4m)
Leaver (=1, Remainer = 0)	0.615	0.084	0.133	0.025	0.412	0.099	-0.219	-0.042
	(0.576)	(0.080)	(0.266)	(0.052)	(0.302)	(0.073)	(0.195)	(0.037)
Norms (1 = majority socially appropriate)								
bribe1 (Leaver)	-1.691***	-0.219***	-1.244**	-0.229***	-0.044	-0.011	0.001	0.000
	(0.608)	(0.075)	(0.511)	(0.088)	(0.529)	(0.126)	(0.461)	(0.089)
bribe1 (Remainer)	-6.142***	-0.794***	-0.201	-0.037	6.473***	1.544***	-0.296	-0.057
	(1.544)	(0.159)	(0.641)	(0.116)	(0.506)	(0.145)	(0.423)	(0.080)
bribe1 (Own)	8.527***	1.102***	1.937**	0.357**	-6.293***	-1.501***	0.400**	0.077**
	(1.875)	(0.183)	(0.967)	(0.164)	(0.745)	(0.193)	(0.203)	(0.039)
bribe2 (Leaver)	1.141**	0.147**	-1.785**	-0.329***	-6.962***	-1.660***	0.260	0.050
	(0.492)	(0.058)	(0.701)	(0.127)	(0.360)	(0.123)	(0.365)	(0.070)
bribe2 (Remainer)	1.478	0.191	-4.214***	-0.777***	1.062***	0.253***	-0.590***	-0.114***
	(1.053)	(0.125)	(1.030)	(0.226)	(0.276)	(0.072)	(0.204)	(0.040)
bribe2 (Own)	-3.864**	-0.500***	5.036***	0.929***	5.211***	1.243***	-0.254	-0.049
	(1.790)	(0.193)	(0.883)	(0.206)	(0.648)	(0.169)	(0.488)	(0.093)
govtrust	-0.133*	-0.017*	-0.007	-0.001	0.015	0.004	-0.001	-0.000
	(0.075)	(0.009)	(0.048)	(0.009)	(0.040)	(0.009)	(0.023)	(0.004)
govtrust19	-0.092	-0.012	0.003	0.000	0.069	0.016	0.163***	0.031***
	(0.081)	(0.010)	(0.046)	(0.009)	(0.055)	(0.013)	(0.037)	(0.007)

Continued on next page

Table 4.7 – continued from previous page

Dep. variable: Reporting positive payoff	Standard		Modified (Leaver)		Modified (Remainer)		Modified (British)	
Model	(1)	(1m)	(2)	(2m)	(3)	(3m)	(4)	(4m)
proud of vote	-0.232***	-0.030***	-0.179**	-0.033**	0.010	0.002	-0.032	-0.006
	(0.090)	(0.011)	(0.088)	(0.017)	(0.058)	(0.014)	(0.051)	(0.010)
winner	-0.989*	-0.128*	0.167	0.031	0.065	0.015	-0.186	-0.036
	(0.565)	(0.068)	(0.255)	(0.047)	(0.349)	(0.083)	(0.137)	(0.026)
diffvote	-0.555***	-0.072***	0.012	0.002	-0.220	-0.052	-0.095	-0.018
	(0.204)	(0.025)	(0.243)	(0.045)	(0.181)	(0.044)	(0.113)	(0.022)
align	0.513	0.066	0.026	0.005	-0.673***	-0.160***	-0.228	-0.044
	(0.405)	(0.054)	(0.204)	(0.038)	(0.214)	(0.049)	(0.178)	(0.034)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	330	330	470	470	520	520	410	410

Notes: Column (1) presents Probit regression, and column (2) presents the marginal effects. Robust standard errors clustered at the subject level in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

4.6 Results 2 (Zimbabwe)

In the following section, we report the results of the experiments conducted in Zimbabwe. We follow a similar procedure as used in experiment 1. We began by investigating dishonest behaviour. In addition, we examined the level of dishonesty dependent on the identity of the player and the receiver. Lastly, we investigated the influence of social norms on dishonesty.

Table 4.8: Summary Statistics of experiment conducted in Zimbabwe

	MG	**MG (Nd)**	**MG (Sh)**	**MG (Zim)**
Identity				
Ndebele (%)	30.30	25.37	24.24	20.90
Shona (%)	69.70	74.63	75.76	79.10
Self-identity (%)				
African	43.94	44.24	40.91	44.78
Ndebele	7.58	7.75	7.58	2.99
Shona	3.03	5.22	6.06	11.94
Zimbabwean	45.45	42.79	45.45	40.30
Socio-economic characteristics				
Age (%)				
18-24	60.61	50.00	53.03	50.75
25-34	30.30	35.07	31.82	34.33
35-44	9.09	14.93	13.64	14.93
Female (%)	57.58	64.18	66.67	58.21
Single (%)	71.21	75.37	72.73	71.64
Household size	4.14	4.67	4.70	4.58
Risk seeking	7.73	7.27	7.42	7.10
Faith in gov	3.52	3.87	3.85	3.36
Risk with gov	5.80	5.86	5.86	5.41
Observations	52	54	51	52

Notes: Baseline summary statistics of participant characteristics in the experiment.

4.6.1 Identity and Dishonesty

Figure 4.3 presents the difference in the fraction of participants who reported "Yes" – positive payoff in the standard mind game and the modified mind game, overall, 36.54%. In the first two columns, when the positive payoff goes to the participant in the standard mind game, 38.82% Ndebele participants and 35.43% of Shona participants report a "Yes" on average over the 10 rounds. When the payoff benefits both the player and another Ndebele participant, 32.50% of Ndebele participants and 36.58% of Shona participants report "Yes". In the Mind Game, where the positive payoff goes to the participant and another Shona participant, 27.69% of Ndebele participants and 34.47% of Shona participants report "Yes" when the positive payoff benefits themselves as well as goes to another participant who is Shona. In the final two bars, in the Mind Game, where the positive payoff goes to the participant but also a Zimbabwean participant, 26.15% of Ndebele participants and 26.92% of Shona participants report "Yes" when the positive payoff benefits themselves as well as goes to another participant who is Shona.

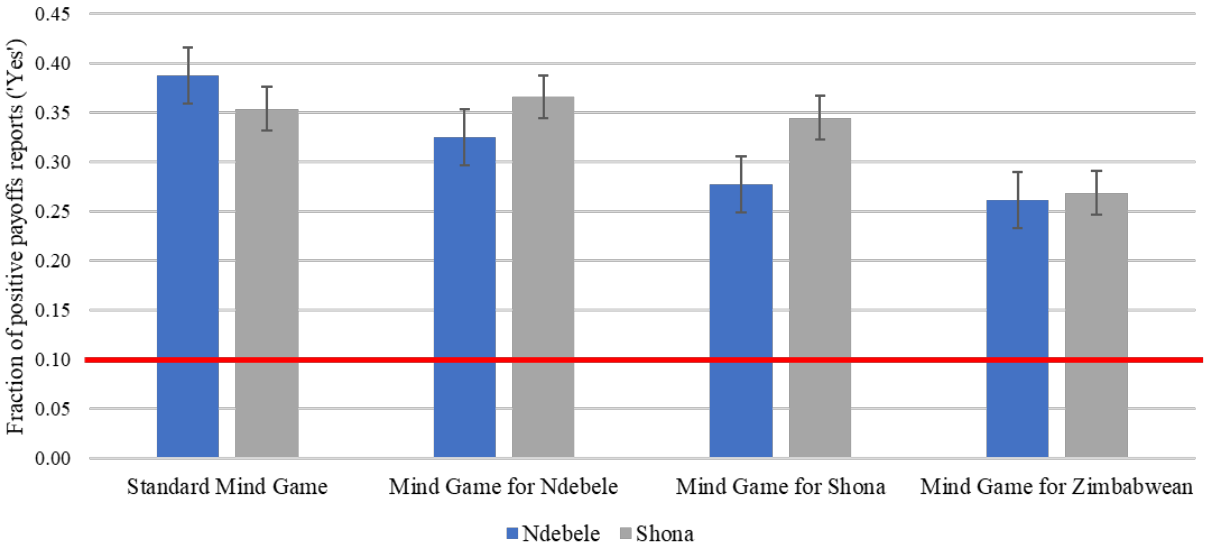


Figure 4.3: Lying the Mind Game - Zimbabwe

Note: Bars present the fraction of 'Yes' outcomes reported. The first set of columns, presents the fraction of 'Yes' outcomes reported in the Standard Mind Game. The red line marks the expected 10% outcome. Error bars indicate 95% confidence intervals.

The expected number of correct guesses (reporting "Yes") is 10%. Again, any value greater than 10% is considered to be due to dishonesty. Participants lie significantly above 10%, across all game variants (binomial test, p=0.000). We find a significant number of reported

"Yes" outcomes across treatment variations by the two identities over the expected 10% (binomial test p =0.000 across all treatments). Confirming Hypothesis 1, we find evidence of dishonest behaviour. We also find evidence of lying aversion as individuals are not willing to report a "Yes" to maximise their payoff in the Standard Mind Game.

Results from non-parametric tests show that in the standard Mind Game, differences in reported "Yes" between Ndebele and Shona participants are not statistically different (p =0.510). In the Mind Game for Ndebele participants, differences in reported "Yes" between Ndebele and Shona participants are not statistically different (p =0.421). In the Mind Game for Shona participants, differences in reported "Yes" between Ndebele and Shona participants are not statistically different (p =0.064). We fail to find evidence that Ndebele and Shona participants report "Yes" outcomes significantly differently in the treatment, where the other receiver is a Zimbabwean player (p =0.982). Therefore, the overall likelihood of dishonesty does not significantly vary between the two identity groups across treatments, except for when the payoff goes to a Shona receiver.

We run Probit regressions with dishonesty as the dependent variable, where the base category is the control group of Ndebele participants, including controls for the demographic characteristics of participants. We also report marginal effects from the Probit model [10]. We find that the dummy variable *Shona* is not statistically significant across all treatments. However, qualitatively, we observe a negative coefficient when the other receiver is Ndebele, perhaps suggestive of spiteful honesty. However, we find that lying on behalf of others may be driven by other factors related to identity formation. Participants were allowed to further self-identify as either African, Ndebele, Shona or Zimbabwean. Using the marginal effects, we find that participants who chose to be Shona report a positive outcome of 35.7 percentage points more than those who self-identify as African in treatment 1 when the payoff goes to the player solely. However, we find that those that chose to identify as Shona report lower outcomes when the other receiver is Shona, by 17.9 percentage points. We observe similar patterns for further identification as Ndebele, although the results are insignificant. Conversely, if an individual self-identifies as Zimbabwean, the likelihood of dishonest behaviour increases by 13.5 percentage points. Therefore, we find evidence of image concerns or moral costs when the other receiver is the in-group. However, dishonest behaviour may be bounded not by ethnic identities but by national identities. Those individuals who identify as Zimbabwean report a higher fraction of "Yes" outcomes for a Zimbabwean receiver. The results are unchanged using random-effects, see Appendix C Table C2.

[10] Observations are dropped from the analysis due to missing data in some of the control variables

Table 4.9: Probit regression of dishonesty with margins (Zim)

Dep. variable: Reporting positive payoff	Standard		Modified (Nd)		Modified (Sh)		Modified (Zim)	
Model	(1)	(1m)	(2)	(2m)	(3)	(3m)	(4)	(4m)
Shona (=1, Ndebele=0)	0.280	0.101	-0.130	-0.048	0.454	0.146	0.086	0.026
	(0.277)	(0.098)	(0.365)	(0.137)	(0.423)	(0.126)	(0.306)	(0.091)
Self-identity (base = African)								
Ndebele	0.615	0.240	-0.489	-0.153	0.043	0.017	-0.612	-0.111*
	(0.407)	(0.157)	(0.440)	(0.116)	(0.369)	(0.142)	(0.446)	(0.062)
Shona	0.928***	0.357***	0.108	0.040	-0.528*	-0.179**	1.444***	0.518***
	(0.301)	(0.107)	(0.351)	(0.131)	(0.276)	(0.084)	(0.319)	(0.107)
Zimbabwean	-0.087	-0.031	0.112	0.041	-0.497*	-0.170*	0.443**	0.135**
	(0.251)	(0.089)	(0.217)	(0.080)	(0.301)	(0.097)	(0.187)	(0.057)
Shona align	-0.022	-0.008	0.027	0.010	0.112	0.039	-0.134	-0.041
	(0.137)	(0.051)	(0.130)	(0.048)	(0.210)	(0.072)	(0.098)	(0.030)
Ndebele align	0.139	0.051	0.096	0.035	0.267	0.092	0.229	0.071
	(0.109)	(0.041)	(0.158)	(0.058)	(0.165)	(0.057)	(0.154)	(0.047)
govtrust	-0.016	-0.006	0.020	0.007	0.055	0.019	-0.002	-0.001
	(0.044)	(0.016)	(0.062)	(0.023)	(0.047)	(0.017)	(0.049)	(0.015)
govrisk	0.078	0.029	-0.014	-0.005	-0.018	-0.006	0.060	0.018
	(0.063)	(0.023)	(0.048)	(0.018)	(0.047)	(0.016)	(0.046)	(0.014)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-0.205		-2.276**		-4.102***		-1.508**	
	(0.931)		(0.937)		(1.403)		(0.735)	

Continued on next page

Table 4.9 – continued from previous page

Dep. variable: Reporting positive payoff	Standard		Modified (Nd)		Modified (Sh)		Modified (Zim)	
Model	(1)	(1m)	(2)	(2m)	(3)	(3m)	(4)	(4m)
Observations	500	500	510	510	490	490	470	470

Notes: Column (1) presents Probit regression, and column (2) presents the marginal effects. Robust standard errors clustered at the subject level in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

4.6.2 Social Preferences and dishonesty

All participants took part in the social norms elicitation task. We take a similar approach to Krupka & Weber (2013). To begin, we converted participants' responses into numerical scores. Participants who evaluate an action to be *very socially inappropriate* is valued at -1, *somewhat socially inappropriate* is valued at -1/3, *Somewhat socially appropriate* is valued at 1/3 and *very socially appropriate* is valued at 1. Tables 6 and 7 present participants' evaluations of social appropriateness. We represent the findings as follows; their perceptions, the Shona majority or the Ndebele majority, for the entire sample of participants. We also identify "if the participant is Ndebele" and "if the participant is Shona".

Table 4.10: Reporting Bribery (Zimbabwe)

	−−	−	+	++	N
Own perception	40.35	12.28	11.84	35.53	225
Ndebele majority	30.40	20.26	22.91	26.43	225
Shona majority	33.62	20.96	23.58	21.83	225
Participant is Ndebele					
Own perception	42.37	6.78	11.86	38.98	59
Ndebele majority	40.68	11.86	22.03	25.42	59
Shona majority	32.20	20.34	18.64	28.81	59
Participant is Shona					
Own perception	40.12	14.37	10.78	34.73	166
Ndebele majority	27.11	23.49	22.29	27.11	166
Shona majority	34.52	21.43	24.40	19.64	166

Notes: −− for ratings `"Very socially inappropriate"`, − for ratings `"Somewhat socially inappropriate"`, + for ratings `"Somewhat socially appropriate"`, and ++ for ratings `"Very socially appropriate"`. Modal responses are shaded. N is the number of observations.

Table 4.10 presents the reported perceptions of socially appropriate actions in the scenario where bribery is reported. Overall, reporting bribery is seen as *very socially appropriate*. We compare the differences in social norms regarding reporting bribery between Ndebele and Shona participants. We find no significant differences in an individual's perceptions when

Table 4.11: Punishing Bribery (Zimbabwe)

	−−	−	+	++	N
Own perception	4.44	5.33	26.67	63.56	222
Ndebele majority	7.62	13.45	32.39	46.64	222
Shona majority	6.67	17.33	29.78	46.22	222
Participant is Ndebele					
Own perception	3.39	6.78	25.42	64.41	59
Ndebele majority	10.17	10.17	22.03	57.63	59
Shona majority	8.47	16.95	27.12	47.46	59
Participant is Shona					
Own perception	4.85	4.85	27.27	63.03	163
Ndebele majority	6.75	14.72	36.20	42.33	163
Shona majority	6.06	17.58	30.91	45.45	163

Notes: −− for ratings `"Very socially inappropriate"`, − for ratings `"Somewhat socially inappropriate"`, + for ratings `"Somewhat socially appropriate"`, and ++ for ratings `"Very socially appropriate"`. Modal responses are shaded. N is the number of observations.

the participant is Ndebele or Shona (Wilcoxon rank-sum test p=0.430). The perceptions of most Ndebele and Shona people are statistically different between Ndebele and Shona participants (rank-sum test p=0.001, p=0.010, respectively). Therefore, we find evidence of diverging norm perceptions between the two identity groups.

Table 4.11 presents the evaluations of the social appropriateness of punishing bribery. Punishing bribery is generally evaluated as being *very socially appropriate*. We compare the differences in social norms regarding punishing bribery between Ndebele and Shona participants. We find no significant differences in an individual's perceptions (Wilcoxon rank-sum test p=0.601). The perceptions of most Ndebele people are statistically different between Ndebele and Shona participants (p=0.000). Most Shona people's perceptions are not statistically different between Ndebele and Shona participants (p=0.972). These results suggest that the distributions of evaluations do not differ between Ndebele and Shona participants, except when evaluating the majority of Ndebele participants' view of punishing bribery.

Table 4.12, reports Probit regression model on reported `"Yes"` outcomes inclusive of social norm evaluations. We ensure robustness by including a set of controls, such as age, gender, risk-seeking, employment status and household size. We hypothesised that dishonesty would be a function of the prevailing social norms. Hence, modal evaluations will impact the propensity for dishonest behaviour. The social norms evaluations revealed that reporting

bribery was socially inappropriate. However, punishing bribery is *very socially appropriate*. We anticipate finding norm conformity; if the majority disapprove of dishonesty or bribery, then there will be a lower fraction of `"Yes"` reports. Surprisingly, our results show that normative evaluations of reporting bribery, *Bribe1 (Sh)* for most Shona people, increase the likelihood of dishonesty when the other receiver is Shona.

Conversely, the evaluations of most Ndebele people's perceptions of punishing bribery, *Bribe 2 (Nd)* decreases the likelihood of dishonesty in the Standard Mind Game. The evaluations of most Shona people's perceptions of punishing bribery decrease the likelihood of dishonesty when the other receiver is Ndebele but increases dishonesty in the Standard Mind Game. In line with our theoretical model, the social element of social norms is observed in our results, as the impacts of norms are primarily noticed in the treatments with another player[11].

[11]See Appendix C, Table C4 for further robustness checks using Poisson regression models. This does not change the results. In addition, we replicate the set of regressions using random-effects model to account for changes in behaviour over the 10 rounds. The results are unchanged. See Appendix C, Table C2

Table 4.12: Probit regression of dishonesty and social norms (Zimbabwe)

Dep. variable: Reporting positive payoff	Standard		Modified (Nd)		Modified (Sh)		Modified (Zim)	
Model	(1)	(1m)	(2)	(2m)	(3)	(3m)	(4)	(4m)
Shona	0.249	0.089	-0.189	-0.070	0.222	0.074	0.156	0.046
	(0.306)	(0.108)	(0.386)	(0.144)	(0.494)	(0.159)	(0.338)	(0.098)
Norms (1 = majority socially appropriate)								
Bribe1 (Nd)	-0.165	-0.061	-0.344	-0.122	-0.170	-0.058	0.120	0.037
	(0.326)	(0.120)	(0.265)	(0.093)	(0.259)	(0.087)	(0.211)	(0.065)
Bribe1 (Sh)	-0.273	-0.099	0.228	0.083	0.738*	0.259*	0.017	0.005
	(0.269)	(0.098)	(0.249)	(0.091)	(0.447)	(0.156)	(0.253)	(0.078)
Bribe1 (Own)	0.451	0.167	0.227	0.083	-0.386	-0.131	-0.272	-0.084
	(0.288)	(0.107)	(0.258)	(0.094)	(0.312)	(0.107)	(0.235)	(0.073)
Bribe2 (Nd)	-1.335***	-0.496***	-0.060	-0.022	-0.153	-0.054	0.388	0.106
	(0.443)	(0.142)	(0.307)	(0.113)	(0.356)	(0.126)	(0.271)	(0.065)
Bribe2 (Sh)	0.555*	0.191**	-0.888***	-0.335***	-0.036	-0.012	-0.365	-0.120
	(0.290)	(0.094)	(0.338)	(0.124)	(0.420)	(0.146)	(0.308)	(0.109)
Bribe2 (Own)	0.368	0.127	0.526	0.172*	-0.324	-0.119	-0.514	-0.177
	(0.349)	(0.110)	(0.331)	(0.094)	(0.478)	(0.184)	(0.330)	(0.123)
govtrust	-0.043	-0.016	0.068	0.025	0.044	0.015	-0.029	-0.009
	(0.049)	(0.018)	(0.075)	(0.027)	(0.049)	(0.017)	(0.059)	(0.018)
govrisk	0.041	0.015	-0.046	-0.017	0.013	0.004	0.086	0.027
	(0.067)	(0.025)	(0.050)	(0.018)	(0.053)	(0.018)	(0.054)	(0.016)
Self-identify (Base = African)								

Continued on next page

Table 4.12 – continued from previous page

Dep. variable: Reporting positive payoff	Standard		Modified (Nd)		Modified (Sh)		Modified (Zim)	
Model	(1)	(1m)	(2)	(2m)	(3)	(3m)	(4)	(4m)
Ndebele	0.192	0.076	-0.429	-0.133	0.239	0.093	-0.424	-0.087
	(0.406)	(0.161)	(0.462)	(0.125)	(0.348)	(0.137)	(0.483)	(0.080)
Shona	0.710**	0.275**	0.110	0.040	-0.814*	-0.248**	1.512***	0.543***
	(0.339)	(0.124)	(0.445)	(0.164)	(0.467)	(0.103)	(0.383)	(0.126)
Zimbabwean	-0.557**	-0.195**	0.151	0.055	-0.506*	-0.171*	0.368*	0.110
	(0.272)	(0.094)	(0.216)	(0.080)	(0.296)	(0.094)	(0.219)	(0.068)
shona align	-0.023	-0.009	0.038	0.014	0.154	0.053	-0.126	-0.039
	(0.146)	(0.054)	(0.124)	(0.045)	(0.216)	(0.074)	(0.094)	(0.029)
ndebele align	0.217	0.080	-0.008	-0.003	0.301*	0.103*	0.279	0.085
	(0.141)	(0.052)	(0.173)	(0.063)	(0.164)	(0.057)	(0.178)	(0.055)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-0.733		-1.890*		-4.188***		-1.239	
	(1.004)		(0.986)		(1.372)		(0.870)	
Observations	500	500	510	510	490	490	470	470

Notes: Robust standard errors clustered at subject level in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

4.7 Discussion and conclusion

Dishonesty and lying for others can have several economic implications. For example, Erat & Gneezy (2012) notes the positive spillovers of a manager lying to an employee that yields greater performance after the fact. However, lying to benefit oneself and others can have negative and damaging consequences as it can lead to corruption (Gächter & Fehr 1999, Benistant & Villeval 2019, Kocher et al. 2018, Weisel & Shalvi 2015). In this paper, we set out to investigate the propensity for dishonest behaviour when the individual benefits from it and when someone else in the in-group or out-group member can also benefit from it. We hypothesise that affinity with the other player will lead to higher reported `"Yes"` outcomes. The research uses two naturally-occurring identity groups in the United Kingdom and Zimbabwe. We compared political groups in the United Kingdom and ethnic groups in Zimbabwe.

Consistent with the theory, we find evidence of dishonest behaviour. However, only partially, thus there is a cost of cheating behaviours (Fischbacher & Föllmi-Heusi 2013, Mazar et al. 2008, Kajackaite 2018). Material payoffs can likely diminish or lessen moral considerations around cheating (Bowles & Polania-Reyes 2012). Nonetheless, participants viewed dishonest behaviour as morally incorrect after a certain threshold (Gneezy et al. 2018). Therefore, a purely rational utility maximisation theory fails to predict dishonest behaviour; instead, the findings are consistent with a partial aversion to lying (Gneezy et al. 2018).

We hypothesised that when others benefit from dishonesty, this changes the likelihood of dishonest behaviour. Our results are consistent with Fischbacher & Föllmi-Heusi (2013) and Benistant & Villeval (2019); the participant's identity did not significantly alter dishonesty when another benefited from it. However, in both experiments, participants' attachment and self-categorisation were the main factors in the likelihood of lying for others. Specifically, in the UK experiment, the relevance of the identity is overshadowed by the feelings towards the identity. For example, a sense of pride and feeling like a winner associated with their voting decision decreases dishonest behaviours. The results from the self-identification in Zimbabwe similarly suggest a moral cost of dishonest behaviour, further affinity to a group reduced dishonesty behaviour. However, the results revealed that dishonesty is bounded by national, not ethnic identity; identifying as Zimbabwean increased dishonesty for Zimbabwean receivers. Therefore, the findings complement Weisel & Shalvi (2015), as we identify that the salience and association of natural identities are more likely to influence dishonest behaviours when dishonesty benefits others in paired environments. In addition, the findings

share similarities with Chowdhury et al. (2016), who find that natural rather than artificial group identities impacted group behaviours. Hence, these findings further confirm results from chapter 2 and chapter 3, where the strength of the group affiliation influences not only prosocial behaviours, but also dishonesty behaviours.

We also explicitly tested the impact of social norms around dishonest behaviour, specifically reporting and punishing bribery. The social norms tasks allow us to identify some justification for dishonest behaviour. Social norms framed around reporting and punishing bribery impacted dishonesty in the UK. In general, we find evidence of norm conformity to the majority's response to bribery, albeit not always consistent. Our findings are consistent with Bicchieri et al. (2021), who report that when it suits individuals, they will comply with the norm, but at other times they will distort their beliefs. In both experiments, we observe differences in behaviour in different norm environments, reporting or punishing bribery. Therefore, we find varying degrees of norm-dependent behaviour.

Briefly commenting on the qualitative differences between the two countries where the experiments were conducted. We observed that, on average, a greater fraction of `"Yes"` reported in Zimbabwe than in the UK. This is contrary to Abeler et al. (2019), who find that dishonest behaviour is similar across countries. Notably, it is often argued that developing countries face higher levels of corruption; therefore, perceptions of honesty and cheating may alter and impact behaviour (Gächter & Schulz 2016, Di Zheng et al. 2020, Hugh-Jones 2016). Thus, our results align with findings by Gächter & Schulz (2016), who find that participants in the United Kingdom lied relatively less than other European, African and Asian countries in the die-roll task. The low levels of dishonesty observed in the United Kingdom may be due to incentives not being a significant enough temptation, as some research has shown that increasing the reward size the likelihood of dishonest behaviour (Gerlach et al. 2019, Gneezy 2005, Kajackaite & Gneezy 2017). On the other hand, perhaps the environment of political dishonesty in developing countries becomes socially acceptable, meaning the society does not deem lying for themselves and others as morally incorrect and thus does not threaten social image concerns (Gächter & Schulz 2016, Dai et al. 2018, Khalmetski & Sliwka 2019).

Furthermore, our social vignette provides evidence of the social acceptance of bribery in Zimbabwe. The results show that reporting bribery is seen as *very socially inappropriate* in Zimbabwe. In contrast, the sample of British participants evaluated reporting bribery as being *very socially appropriate*. However, punishing is viewed as *very socially appropriate* in both countries. Therefore, we observe some differences in perceptions surrounding responses to dishonest behaviour. Another finding of interest is the beliefs observed in chapter 2 and

lying in the Mind Game. In chapter 2, Ndebele participants are trusted less, and we find evidence of statistical discrimination that Ndebele participants are likely to contribute less of their points to others.

Interestingly, we observe that Ndebele participants have a higher likelihood of dishonesty in the Standard Mind Game, in absolute terms. Nonetheless, we find little evidence of norm compliance in the second experiment conducted in Zimbabwe. Therefore, our finding is contrary to the research that states that dishonesty in developing countries arises because of overall greater norm acceptance (Di Zheng et al. 2020, Ryvkin et al. 2017, Gächter & Schulz 2016). Although, we are unable to assess the implications of wider institutions on dishonesty.

Our results must be interpreted with caution due to the sample size. In light of this, we use appropriate statistical techniques. Our findings are often consistent with the literature by observing partial dishonesty. Although the paper suffers from a small sample size, we are able to draw some meaning from our results because of this consistency. For instance, although not our main results of interest, consistent with the literature, for instance, we observe that men are more dishonest than females (Gerlach et al. 2019).

In this paper, we show that group identity and the salience of the group identity impact dishonesty by increasing the moral cost of dishonest behaviour. Our results show that social norms around bribery shape dishonest behaviour. Future research can further understand these mechanisms and utilise these to reduce corruption (Jiang 2013). We find some evidence that the more individuals view reporting and punishing bribery as socially appropriate, the less dishonesty is observed in the mind games. This suggests that a bottom-up approach based on emphasising the social norms around bribery may effectively reduce dishonest behaviour and, thus, bribery. Our results also highlighted the importance of trust in institutions. Government trust and faith in government significantly impacted dishonesty on behalf of others.

Our result is in line with findings by Druckman et al. (2021), who suggests that polarisation along political lines impacts perceptions of COVID-19 responses and policies. Our findings support the argument that rather than being a developing country issue, it is due to the "Prevalence of rule violations" (Gächter & Schulz 2016, p.2). Therefore, future research is needed to understand the impacts of government trust, especially following COVID-19 in the United Kingdom. Another extension could further investigate the role of government scandals and dishonesty on citizens' prevailing acceptance of dishonest behaviours.

Future research can also consider the impact of detection in dynamic games. Our present study does not include an exogenous cost associated with lying, such as the element of feedback, retaliation or detection. Previous research has shown that when there is the opportunity for penalties or detection, there are higher levels of honesty than traditionally predicted by the model (Fischbacher & Föllmi-Heusi 2013, Kajackaite & Gneezy 2017). However, investigating dishonesty behaviour when there is a possibility of detection can be used to understand further willingness to lie when there is greater risk and also to understand the willingness to whistle blow or report unethical behaviour in our context of engaging with and across political lines.

5

Voting outcomes and the Big Five personalities

Abstract

We investigate the relationship between the Big Five personality traits and behaviour. Individual personality traits may influence an individual's decision to leave or remain in the EU. We use the short Big Five Inventory (BFI-S) to assess the role of personality traits in influencing cooperatives, coordination and voting decisions. We look specifically at the five basic personality traits: Neuroticism, Extraversion, Openness to experience, Agreeableness and Conscientiousness. Our results suggest that Leavers and Remainers differ in Openness, with Remainers reporting higher levels of Openness. In addition, we find evidence that personality traits dictate cooperative behaviours. Notably, in further analysis, we also find that conscientiousness and extraversion play a role in a participant's inconsistent reporting of their identity as a Leaver or Remainer, respectively, to the recruitment platform. We provide potential rationale for this finding in the discussion. The findings contribute to the research on personality traits, political decisions and behaviours.

Keywords: Brexit; Cooperative Behaviour; Personality; BFI-S; Leavers; Remainers; Misreporting

5.1 Introduction

The Big Five personality traits provide a schema to investigate individual personality differences and their association with certain behaviours and outcomes, from mobile phone addiction and participation in group action to academic dishonesty (Hong et al. 2012, Robalo et al. 2017, Giluk & Postlethwaite 2015). These personalities are grouped into 5 dimensions: *Extroversion*, *Agreeableness*, *Emotional Stability*, *Conscientiousness* and *Openness* (see John et al. (2008) for a review). The existing literature has also investigated the relationship between political behaviours and individual-level personality traits (Gerber et al. 2013, Hirsh et al. 2010). Personality traits may be a core predictor for British voters who decided to vote either Leave or Remain in the 2016 European Union (EU) referendum and the propensity for cooperative behaviour. Therefore, using controlled online experiments, this paper investigates whether individual personality traits explain their decisions in the EU referendum and the degree to which personality traits can explain inter-group behaviours between Leavers and Remainers.

Research into the characteristics of Leave and Remain voters conducted by Alabrese et al. (2019) identified different individual-level attributes closely associated with each voting individual's decision, such as ethnicity, education and life satisfaction[1]. In addition, Boomgaarden et al. (2011) identifies five dimensions of Euroscepticism. *Euroscepticism* can be due to attitudes and emotions that fear what the EU could bring about. The second aspect concerns the importance of European identity. Thirdly, Euroscepticism can be based on institutional concerns. The fourth consideration is the utilitarian benefit assessment of EU integration. Lastly, it concerns future integration with the EU and long-term policies and relationships.

Similarly, Hobolt (2016) suggests that behind Euroscepticism are concerns about identity. Moreover, Euroscepticism has been identified as a general fear or enmity towards other cultures, such as minority groups and immigrants. These opinions are mainly formed based on different cues (Anderson & Mellor 2009). Focusing on the first and second dimensions of Euroscepticism is associated with hostility towards other cultures and fear of integration. These attitudes and emotions are mirrored in Big Five personality descriptors related to fear and hesitancy to engage with new and different environments, contexts and people - such as Neuroticism and low levels of Openness. To the best of our knowledge, we are the first to examine if these attitudes and views that culminated in the decision to either vote to leave

[1]See chapter 3 for further background into the 2016 EU Referendum and identity formation.

or remain in the EU are contingent on individual-level personality characteristics. Here, we add to our analysis in chapter 3 by considering the individual-level personality traits at play when interacting with Leavers, Remainers, Non-voters and British players.

Seminal research by Goldberg (1992) identified prototypical traits that consistently produce the Big Five dimensions, *Extroversion*, *Agreeableness*, *Emotional Stability*, *Conscientiousness* and *Openness*. The use of the psycho-lexical method - understanding individual-level personality traits in language - identified five broad personality traits in the English language (Goldberg 1992, 1993)[2]. These personality traits identify wide-ranging and expansive psychological characteristics (John et al. 2008, Gerber et al. 2013). Over several studies, the authors used a self-reflective questionnaire combining unipolar, bipolar and adjective questions to identify five ex-post broad personality traits. The findings showed that meaningful and generalisable personality traits could be successfully identified using the Big Five dimensions. Since then, there has been extensive research on the Big Five personality traits. The psychological literature has expanded on this and is able to derive the same personality traits in different languages (McCrae & Costa 1997). A shorter Big Five personality measure has been utilised in psychological and economic research Rammstedt & John (2007), Gerlitz & Schupp (2005), highly correlated with the full BFI, allowing the experimenter to quickly and effectively assess personality traits (Hahn et al. 2012). In this paper, we measure personality characteristics using the short BFI (BFI-S).

Several studies have investigated the relationship between the Big Five personality traits and political beliefs. They have been shown to influence voter turnout and voting patterns. Graham et al. (2009) find that Liberals and Conservatives differ in their moral foundations. Openness and Conscientiousness have been found to predict voting patterns for Democratic and Republican runners in the U.S, respectively. (Rentfrow et al. 2009). Similarly, research conducted in Italy using a shortened version of the BFI show that individuals high in Openness and friendliness are more likely to support centre-left programs (Caprara & Schwartz 2006). Further, Hirsh et al. (2010) observe that facets of Agreeableness are associated with Liberalism, and aspects of Conscientiousness were more closely associated with Conservatives. The research has shown a positive relationship between Left-leaning politics and Openness (Gerber et al. 2013, Hirsh et al. 2010, Rentfrow et al. 2009, Jost et al. 2003).

A growing literature maps economic concepts with psychological views of personalities (see, Thielmann et al. (2020) for a review). The underlying motivation for cooperative behaviour may be due to psychological characteristics (Swope et al. 2008, Rustichini et al. 2016, Hilbig

[2]Lang et al. (2011) for analysis of different assessment methods of the Big Five dimensions.

et al. 2012). Collective behaviour has been identified as being partially dependent on the identity of those coming together in a cooperative setting (Akerlof & Kranton 2000, 2005), even their political identity (Klein et al. 2019). Although other cultural and social factors such as social norms may influence behaviour (Ostrom 1998, Bénabou & Tirole 2011*a*). All aspects of the Big Five personality traits have proven to have high explanatory power in cooperative behaviour (Ben-Ner et al. 2004, Boone et al. 1999, Swope et al. 2008, Brandstätter & Güth 2002). Hence, the addition of individual-level personality traits extends the experimentalist's ability to make generalisable findings from experiments. Ben-Ner & Kramer (2011) find that personality traits impact behaviour. Specifically, the authors find that personality traits influenced giving in the dictator game; extraversion negatively affects giving by male participants and neuroticism negatively effected giving by female participants. These gender differences are consistent with gender differences observed across economic games (Eckel & Grossman 2008, Slonim & Guillen 2010, Iriberri & Rey-Biel 2021).

Furthermore, when investigating identities such as those formed following the Referendum. There is the danger that participants will misstate their identity, whether intentionally, unintentionally, or due to experimenter demand effects - changes in behaviour or reporting by participants towards what they perceive experimenter deems to be appropriate or correct behaviour (Barber & Silver 1968). Subjects misreporting in experiments have been widely researched from clinical to accounting research (Achilles et al. 2018, Murphy 2012). As more research is conducted online, a growing avenue of research is investigating issues of low data quality due to misrepresentation, low attention and the potential for bots (Chandler & Paolacci 2017, Wessling et al. 2017, Chmielewski & Kucker 2020, Hydock 2018). We, therefore, also test the validity of online participant recruitment services, in our case - Prolific Academic. When using online recruitment services for experiments, participants reported identity must be consistent, especially when investigating discriminatory behaviour. When participants sign up to the platform, they are asked to provide information about themselves that can be used in screening for research. Participants are incentivised to answer these additional characteristics questions as it means they will be eligible to partake in a greater number of studies (Palan & Schitter 2018). The recruitment process of participants should be an essential consideration in running behavioural economic experiments[3].

Attitudes and personalities that are more inclined to misreport can have implications on conducting trials and fraud. To date, the majority of economic and psychological studies have focused on academic dishonesty and personality traits (DeAndrea et al. 2009, Mark

[3]See, for example, Greiner (2015) for a list of reasons why experimenters should be aware of the importance of the recruitment process.

& Ganzach 2014, Barthelemy & Lounsbury 2009). Murphy (2012) find that when given a chance, higher Machiavellians - cynical individuals that do not trust others and are willing to be manipulative (Burks & Krupka 2012, p.206) - misstate without feeling high levels of guilt. Oh et al. (2011) find that subjects with high extraversion and low honesty-humility were more likely to show signs of workplace deviance. Similarly, MacInnis et al. (2020) find that subjects with lower honesty-humility characteristics are more likely to misrepresent their age. Given the nature of the identities, this paper also focuses on - whether willfully or not - incorrect reporting of identities can have significant implications. In this paper, we also compare the identity reported by participants in the experiment about their voting decision in the 2016 EU referendum and their reported identity to Prolific Academic. We explore if certain personality types are more likely to be associated with misreporting. Here we contribute to the literature on the reliability of online research, specifically on identities. Moreover, we are able to glean an understanding of the characteristics of those who misrepresent their identity potentially out of regret.

We examine the differences between personality measures and behaviour observed in games. This paper explores cooperative behaviour as expressed by mean contributions in popular experimental games the *prisoner's dilemma*, *dictator game*, *trust game*, *public goods game* and a modified version of the public goods game *(mPGG)*, in relation to psychological characteristics. We investigate how differences in cooperative behaviour, as displayed in economic games, are related to individual-level personality traits. We find that Leavers and Remainers present similar personality characteristics across the Big Five personality traits, except for Openness. Remainers are more open to experience. Secondly, the results provide evidence that personality traits explain cooperative behaviours. Lastly, we find evidence of misreporting; Conscientious Leavers are more likely to misstate their voting decision.

Our contribution is threefold: we contribute to the growing literature around some of the characteristics that led to and influenced voting decisions in the 2016 EU Referendum and the prosocial behaviours of Leavers and Remainers. We also contribute to experimental literature on the validity of online experiments.

The remainder of this paper is organised as follows. In section 5.2, we identify the trait characteristics in relation to political identity and cooperativeness. Then in Section 5.3, we present our experimental design and procedure. Lastly, we discuss the results and conclude our findings.

5.2 Big 5 personality traits, cooperation and voting

In the following section, we discuss the relationship between each Big Five personality trait and how it theoretically and empirically relates to cooperation, altruism, trust and the potential for misrepresentation. We also link the personality traits to socio-demographic characteristics associated with voting to Leave or Remain.

5.2.1 Neuroticism

Neuroticism is associated with the tendency toward anxiety, self-doubt and guilt. Individuals who score as highly neurotic are unable to cope well with stress and are likely to be in denial in response to negative triggers. Psychologists find that high levels of Neuroticism is associated with coordination in the prisoner's dilemma due to fear of repercussions from defecting (Hirsh & Peterson 2009, DeYoung et al. 2007). Conversely, highly neurotic individuals are more likely to experience betrayal and therefore are more hesitant to trust and coordinate with others (Sabater-Grande et al. 2022, Evans & Revelle 2008).

Alabrese et al. (2019), finds that voting to Leave is associated with low life satisfaction and low internet and technology usage. Similarly, the psychological research by Hayes & Joseph (2003) find that Neuroticism and Conscientiousness were closely associated with life satisfaction scores. On the other hand, there have been mixed findings regarding whether there is a positive or negative association between Neuroticism and internet usage. Mark & Ganzach (2014) find that Neuroticism is positively associated with online communication because neurotic individuals may use online communication to prevent feelings of isolation and loneliness. Following the referendum, Ashcroft (2016) also identifies fear and risk surrounding leaving the EU were more associated with Remainers. This in turn can be associated with Neuroticism.

We then hypothesise that Remainers are more likely to score high in Neuroticism. Additionally, we hypothesise that Neurotic individuals will be less cooperative and more likely to display out-group discrimination. However, we also assume that the tendency for self-doubt and guilt will likely lead to consistent identity categorisation between the recruitment platform and the experimenter.

5.2.2 Extraversion

Extraversion is associated with assertiveness and being highly comfortable in social settings. Extraverted individuals are likely to be highly prosocial (Costa & McCrae 1992, Goldberg 1992, DeYoung et al. 2007). In contrast, individuals who are low in extraversion and, therefore, more introverted are more reserved. It is worth noting that Neuroticism and extraversion are not bipolar markers of one another. However, antonyms of Neuroticism are shared with markers of extraversion.

Focusing on the mid-level characteristic of extraversion - enthusiasm - Hirsh & Peterson (2009) find that highly enthusiastic individuals coordinate more in the prisoner's dilemma, as they may be optimistic about future cooperation. Extraversion has been linked with happiness, which in turn has been linked with cooperativeness (Lu & Argyle 1991). Similarly, Swope et al. (2008) and Ben-Ner & Halldorsson (2010) find that extraversion leads to higher amounts sent in the trust game.

As previously highlighted, Alabrese et al. (2019) noted that relatively low internet and smartphone usage is associated with a higher probability of voting to leave the EU. Mark & Ganzach (2014) finds that extraversion is positively associated with internet usage for purposes of communication and leisure. Further, extraverts are more likely to report higher levels of life satisfaction (Pavot et al. 1990). Due to an extraverts tendency for sensation seeking, extraversion has been linked with greater deviant or dishonest behaviour, although there have been mixed findings regarding extraversion and dishonest behaviour (Murphy 2012, DeAndrea et al. 2009, McTernan et al. 2014).

We hypothesise that individuals who score highly as extraverted are more likely to have voted Remain. Likewise, we hypothesise that individuals who report high extraversion are more likely to be cooperative and altruistic and less likely to discriminate. Lastly, we hypothesise that extraverts will be more likely to misstate their identity.

5.2.3 Openness

Individuals who score highly as open to experiences are imaginative, seek new experiences, and are creative. Highly open individuals are also less inclined to conform to the status quo. Openness has often been conflated with intellect (McCrae & Costa 1997, Gerber et al. 2013, DeYoung et al. 2007). Here, we follow the Openness characteristics, which state that

individuals will be open to intellectual curiosity and new ideas, as well as being creative (John et al. 1991, Costa & McCrae 1992). Individuals with low Openness are more likely to be hesitant to consider other ideas and opinions, be traditional, and have limited creativity. Behavioural economic and psychological research finds that Openness is positively associated with trust and trustworthiness (Ben-Ner & Halldorsson 2010, Evans & Revelle 2008). Several explanations have been offered for the connection between Openness and prosocial behaviour. The relationship may be explained by "efficiency concerns" (Lawn et al. 2021, p.4) as open individuals are more aware of and concerned about maximising utility, including group payoffs (Deyoung et al. 2014). Potentially open individuals are more inclined to prosocial behaviour in social dilemma games because they are more accepting of uncertainty (Lawn et al. 2021).

Openness has also been linked with being persuadable (Giluk & Postlethwaite 2015). Hence, highly Open individuals may have been persuaded by the majority media had a pro-Brexit bias (Levy et al. 2016). However, high levels of Openness have been traditionally associated with left or liberal political orientation (Hirsh et al. 2010, Klein et al. 2019). Moreover, Ashcroft (2016) and Hobolt (2016) find that Leavers perceive globalisation as a force for bad. Therefore, we assume that fear of globalisation and others is associated with low Openness. Regarding potentially misrepresenting information, highly open individuals tend to exemplify high creativeness, Gino & Ariely (2012) find that creative individuals can justify their actions, leading to greater unethical behaviour and, thus, a greater tendency of dishonesty.

Although Remainers may be more inclined to believe that leaving the EU will work out, which can be driven by high Openness, we hypothesise that highly open individuals are more likely to have voted to remain in the EU, as they fear neither the Openness implicit in EU membership nor the changes associated with growing globalisation. For Leavers, we assume the fear of integration and globalisation has more substantial explanatory power in their voting decision than hopefulness that leaving the EU will work out. In addition, we hypothesise that highly open individuals are more inclined to misreport.

5.2.4 Agreeableness

Agreeableness is associated with altruism, trust and other-regarding considerations (Goldberg 1992, Costa & McCrae 1992, McCrae & Costa 1997, DeYoung et al. 2007). Individuals with low levels of Agreeableness are more likely to be rude, sceptical of others and cold (Costa & McCrae 1992). Thus, Agreeableness is likely to be associated with higher levels

of unconditional cooperation because of altruism (Ben-Ner & Halldorsson 2010, Kagel & McGee 2014, Kline et al. 2019, Sabater-Grande et al. 2022).

Related to political orientation, Agreeableness is identified as being associated with liberal stances (Hirsh et al. 2010). Liberalism is often linked with age as younger people are more likely to be liberal, whereas conservativism increases with age (Danigelis et al. 2007). In a similar vein, Alabrese et al. (2019) finds that Leavers were more likely to be older.

This personality trait is also negatively associated with competitiveness (Graziano et al. 1997), anti-social behaviour (Miller & Lynam 2001), as well as academic dishonesty (Giluk & Postlethwaite 2015). Shalvi et al. (2011), also observe that competitive personality types are more likely to lie to benefit others than themselves. However, people seek to avoid deception due to agreeable individuals' inclination for fairness could drive this.

We assume that highly agreeable individuals are more likely to have voted to Remain. In addition, we hypothesise that individuals high in Agreeableness are likely to be highly trusting, cooperative and altruistic when interacting with others due to Agreeableness being associated with selflessness and a high significance of moral philosophy (Costa et al. 1991, p.888). Thus, agreeable individuals will be less inclined to discriminate (Parks-Leduc et al. 2015). We also hypothesise that highly Agreeable individuals are less prone to misstate between the recruitment platform and the researcher due to the negative association with competitiveness.

5.2.5 Conscientiousness

Conscientiousness is consistent with characteristics such as planning and organisational skills and complies with social norms and rules (Roberts et al. 2014, DeYoung et al. 2007, John & Srivastava 1999). Moreover, highly conscientious individuals will prepare and research before making decisions (Giluk & Postlethwaite 2015). High levels of conscientiousness has also been linked with "moral scrupulousness and cautiousness" (Costa et al. 1991, p.889), as well as authentic pridefulness (Tracy & Robins 2007). In comparison, individuals with low conscientiousness are careless, unorganised and absent-minded (Burger 2015). Conscientious individuals may be more inclined toward prosocial behaviours when there is a positive outcome and, therefore, more inclined to pursue Pareto optimal strategies. Alternatively, being orderly and well-planned may hinder an individual's willingness to trust others, as they prefer to choose dominant outcomes that ensure maximum individual benefit. A meta-analysis

by Kline et al. (2019) reveals mixed results of conscientiousness and prosocial behaviour.

Alabrese et al. (2019) find that Leave voters were more likely to be unemployed prior to the referendum, which may indicate lower job performance due to carelessness as conscientiousness has been linked with job performance (Shaffer & Postlethwaite 2012). Conscientiousness has also been shown to be a negative predictor of academic cheating or deviant behaviour (Hendy & Montargot 2019, Giluk & Postlethwaite 2015). Hirsh et al. (2010) found, in the US, that Republicans are more Conscientious than Democrats. Likewise, Rentfrow et al. (2009) find that Openness and Conscientiousness predicted political beliefs in the US. In the United Kingdom, Ashcroft (2016) notes that voting Leave is associated with conservatism as 58% of Conservatives and a majority of UKIP supporters voted Leave.

We thus hypothesise that Remain voters are more likely to be high in conscientiousness. Also, highly Conscientious individuals will be less cooperative, as they plan and are aware of the trust required in interacting with others (Evans & Revelle 2008, Ben-Ner & Halldorsson 2010). Lastly, highly conscientious individuals are unlikely to misrepresent their identity as they endeavour to do the "right" thing.

5.3 Experimental Method

The following presents the personality traits measurements and the procedures adopted. The experiments used are designed to investigate prosocial behaviour and discrimination, which repeats methods presented and utilised in chapter 3. We used experimental games commonly used in the literature (Forsythe et al. 1994, Güth et al. 1997, Berg et al. 1995) to measure actual rather than reported behaviours (Baumeister et al. 2007).

5.3.1 The BFI-S

Big Five personality dimensions began with large-scale studies. The literature has commonly used a 44-item Big Five personality traits test (John et al. 1991). This paper uses the short Big Five personality test, a shortened measure of the Big Five dimensions. Specifically, we conducted our studies using the 15-item German version of the Big Five dimensions (Gerlitz & Schupp 2005, Rammstedt & John 2007, Lang et al. 2001, John & Srivastava 1999). Short measures of personality traits have been observed to be reliable and comparative to larger-scale tests (Gosling et al. 2003, Hahn et al. 2012). Our short Big Five personality

measure uses a Five-factor model of Neuroticism, Extraversion, Openness to experience, Agreeableness and Conscientiousness. The 15-item self-reported questionnaire is grouped into groups of 3 that construct our Five-factor model [4]. We use unipolar descriptive markers consisting of a 15-item questionnaire rated from 1 (Strongly disagree) to 5 (Strongly agree). See Appendix D Figure D2 for the Big 5 Inventory questions.

5.3.2 Voting patterns

At the beginning of the study, participants were asked to identify if they chose to [Leave or Remain] in the 2016 European Union Referendum. We also asked participants "to what extent do you align with the reported voting identity". Within our sample, 47.3% Leave voters and 52.7% Remain voters. In the post-experiment questionnaire, participants were also asked to identify what they voted for in the 2019 General Election. Our results show that 29.5% of our sample voted for the Conservative party, 35.9% of participants voted for the Labour Party, 6.8% for the Liberal Democrats, 5.5% for the Green Party, 5.4% for the Scottish National Party, 1.4% for the Brexit Party and 15% of the sample did not vote in the General Election.

The sample showed a range of political preferences of Leavers and Remainers. Of those who reported voting to Leave, 54.0% voted for the Conservative Party, and 23.4% voted for the Labour Party. For participants who voted to Remain in the EU, 50.0%, then went on to vote for the Labour Party in 2019 and 11.3% voted for the Conservative Party and Liberal Democrats, respectively. We find a moderate positive correlation between being a Leaver and voting for the Conservative party in 2019. Conversely, we find a moderately weak correlation between being a Remainer and voting Conservative. We find a weak positive correlation between voting to Remain and Labour. However, we also find a weak negative correlation between voting to Leave and voting for Labour. In sum, we find inconsistent patterns between the referendum and General Election voting decisions. See Appendix D for further supplementary material on the referendum and the General Election.

[4]Lang et al. (2011) find that personality traits are robust across various methods, such as self-reported assessments.

5.3.3 Prisoner's dilemma

Players can choose between two options (strategies) - cooperate or defect. The player's final payoff depends on the other participant's choice. The Pareto dominant strategy is the outcome with the least difference between players and the highest payoff for both players, a preference to cooperate (C, C). However, the dominant strategy is the outcome with the highest individual payoff (D, D), even if there is a significant difference between payoffs between participants.

Participants were given neutral language and shown a table with Red and Blue cards. The cooperative outcome is (B, B), and the purely self-interested outcome is (R, R). Participants are tasked with making incentivized decisions between an in-group, out-group and mixed-group (British) receiver; this indicates if preferences toward cooperation change depending on the social identity of the partner. Differences in cooperation between receivers measures taste-based discrimination. We also elicit participants' beliefs regarding the decisions of others. Therefore, we can identify statistical discrimination [5].

5.3.4 Dictator Game (DG)

The dictator game Roth & Erev (1995), is used to measure taste-based discrimination (Lane 2016, Grosskopf & Pearce 2017). The dictator (Player A) was given 10 points for each receiver (a Leave, Remain and British player). The dictator has to decide how much, if any, to give to the three receivers. If the dictator sends an amount to the receiver, the dictator keeps the remaining amount. We also elicit participants' beliefs regarding their beliefs about the decisions of others. Therefore, we can identify statistical discrimination.

5.3.5 Trust Game (TG)

The trust game by Berg et al. (1995) identifies the degree of trust and reciprocity between participants. In this game, the trustor (Player A) is endowed with 10 points for each trustee

[5]We elicit beliefs after all one-shot games. We then match their beliefs ex post (De Geest & Stranlund 2019, Schotter & Trevino 2014, Gächter & Renner 2010). The literature has shown that beliefs about others' behaviour can affect instances of cooperation and prosocial behaviour (Fischbacher & Gächter 2010, Gächter & Renner 2010, Guala et al. 2013). As participants may expect less cooperative and prosocial behaviour from in-group members (Ferraro & Cummings 2007). Gächter & Renner (2010) noted that beliefs were more accurate when incentivized; for this reason, we incentivized beliefs.

(a Leave, Remain and British player). Each point the trustor sends is multiplied by 3, which the trustee then receives. If the trustor is purely self-interested, they will keep all 10 points for themselves. However, if they choose to cooperate, both players will be better off. However, player A is vulnerable as player B may choose to keep all the multiplied points. Therefore, any amount the trustor sends measures player A's trust. The Trust Game measures taste-based and statistical discrimination, as senders consider the potential for reciprocity. We also elicit participants' beliefs regarding the decisions of others.

5.3.6 Public Goods Game (PGG)

In the public goods game (PGG) by Gächter & Fehr (1999), participants are placed in a group with three other members. Players decide how much of their endowment, if any, to contribute to the common pot (or public good). The total contribution is multiplied by 1.5. The contributions will then be equally redistributed between all four participants: the PGG tests cooperativeness and the degree of prosocial behaviour.

The dominant strategy is the purely self-interested strategy to free-ride. In this strategy, the player will keep all of their endowment (10 points) and potentially benefit from the divided total contribution. Alternatively, the Pareto optimal outcome maximises everyone's payoff when all players contribute all their endowments to the common pot. This, however, is a weak equilibrium because of the risk of others' free-riding. The one-shot public goods game identifies taste-based discrimination through the difference in contribution level between in-group and out-group matching. We also elicit the beliefs, as players decide how much they believe others would send to them.

5.3.7 Modified Public Goods Game (mPGG)

We also use a variation of the Standard public goods game to identify if group size matters and if a critical mass significantly alters altruism and cooperation in a similar decision structure, see Figure 3.4. In the modified public goods game (mPGG), participants will play the game with one other member. Therefore, two players will contribute to the public good. Any amount sent will be multiplied by 1.5, and the total contributions will be divided equally between the two players. The mPGG tests cooperativeness and the degree of prosocial behaviour. We also elicit the beliefs of players as they decide how much they believe others would send to them.

5.3.8 Procedure

The experiments were conducted online using the Prolific recruitment service and online survey software Qualtrics (Qualtrics 2020), over three sessions. At the beginning of the experiment, participants received instructions. See Appendix D for instructions[6]. They were informed that they would receive a fixed participation fee (£2.50) and an additional payoff depending on their decisions. All participants played all of the experimental games to receivers - Leavers, Remainers, and UK citizens. They then completed the post-experimental questionnaire and the BFI-S 15-item survey following these tasks. We use a within-subject design to control for idiosyncrasies in participants (D' Adda et al. 2016). Additionally, to ensure that the heterogeneity between these two homogenous groups of interest is not neglected, data such as sex, education, ethnicity, employment status, household size, marital status and religious affiliation, as well as information such as age and income variables are collected to use as controls in the analysis of the research.

5.4 Results

In this section, we first report the summary statistics about the participants' personality traits in the experiment. Using econometric analysis, we then analyse the impact of personality traits on cooperative behaviours. From that, we investigate the relationship between the Big 5 personality traits and the likelihood of misrepresenting voting decisions.

5.4.1 Personality traits and cooperative behaviour

Figure 5.1 provides the first indication of personality differences between voters. The figure shows the ratings for each personality trait by Leavers and Remainers. Conscientiousness and Neuroticism are relatively similar between the two groups. However, we observe mild differences in ratings for Agreeableness and Extraversion. Notably, the greatest difference is Openness ratings, with Remainers reporting greater Openness than Leavers.

[6]The instructions used were neutral and aimed not to detail the objectives of the experiment.

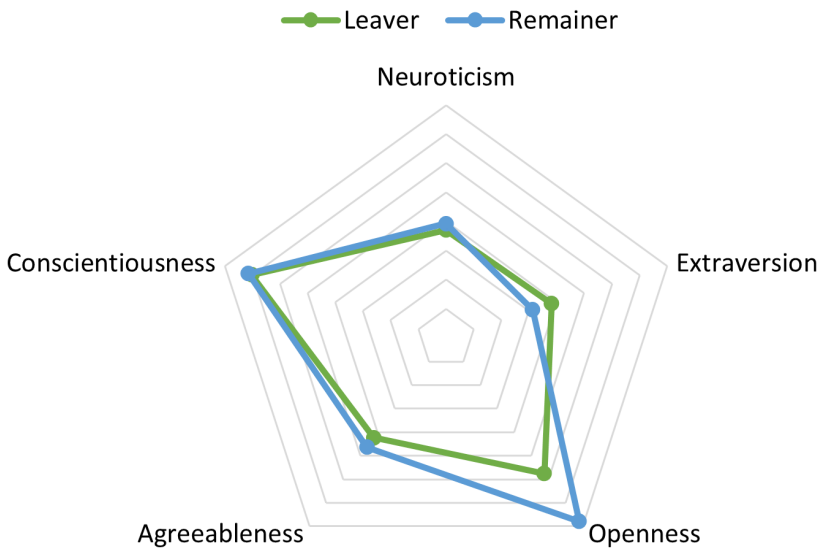


Figure 5.1: Big Five personalities by voting decision

Note: The figure shows a radar chart for the five personality traits. The green radar presents the Leavers personality measure. The blue radar presents Remainers personality measure. The chart is normalised to a scale of 2.8 to 3.8.

We use the BFI-S to identify individual personality traits. Table 5.1 presents summary statistics of the self-reported personality questionnaire for the entire sample and for participants who voted to Leave or Remain in the EU. Overall, on a scale of 1 to 5, participants report Neuroticism, Extraversion and Agreeableness characteristics as being true to some extent (M=3.28, 3.25, 3.34, respectively). Individuals somewhat agree to have Openness and Conscientiousness characteristics (M=3.58, 3.61, respectively). As presented in Table 5.1, Cronbach's alpha for the five scales were as follows: Neuroticism = .49, Extraversion = .54, Openness = .56, Agreeableness = 0.57, and Conscientiousness = .44. We observe low Cronbach's alpha values for our BFI-S. Previous papers have also had low-level alpha values for some indicators (see, for example, (Gosling et al. 2003, Hahn et al. 2012)). Lang et al. (2011) also notes low Cronbach's alpha values for the BFI-S. Nevertheless, when comparing the personality traits with the prevailing literature, our findings are consistent with the literature that Openness predicts political beliefs (Rentfrow et al. 2009).

A similar pattern is observed when we separate the sample between Leavers and Remainers. Figure 5.2 shows the average rating across the five personality traits by Leavers and Remainers. We observe differences in the Big Five characteristics between Leavers and Remainers. The differences in personality traits between the groups are statistically insignif-

icant, however, with the exception of openness, for which we find a statistically significant difference between Leavers and Remainers (Wilcoxon rank-sum test $p = 0.006$; t-test $p = 0.022$). According to these results, an individual's decision to vote to Leave or Remain (in) the European Union is partly due to differences in openness.

Table 5.1: Summary Statistics

	M (SD)	N	E	O	A	C
Neuroticism (N)	3.28 (0.59)	0.49				
Extraversion (E)	3.25 (0.57)		0.54			
Openness (O)	3.58 (0.84)			0.56		
Agreeableness (A)	3.34 (0.55)				0.57	
Conscientiousness (C)	3.61 (0.48)					0.44

Notes: Summary of descriptive variables. N = 505. M, mean; SD, standard deviation. Cronbach's alpha are given in the diagonal.

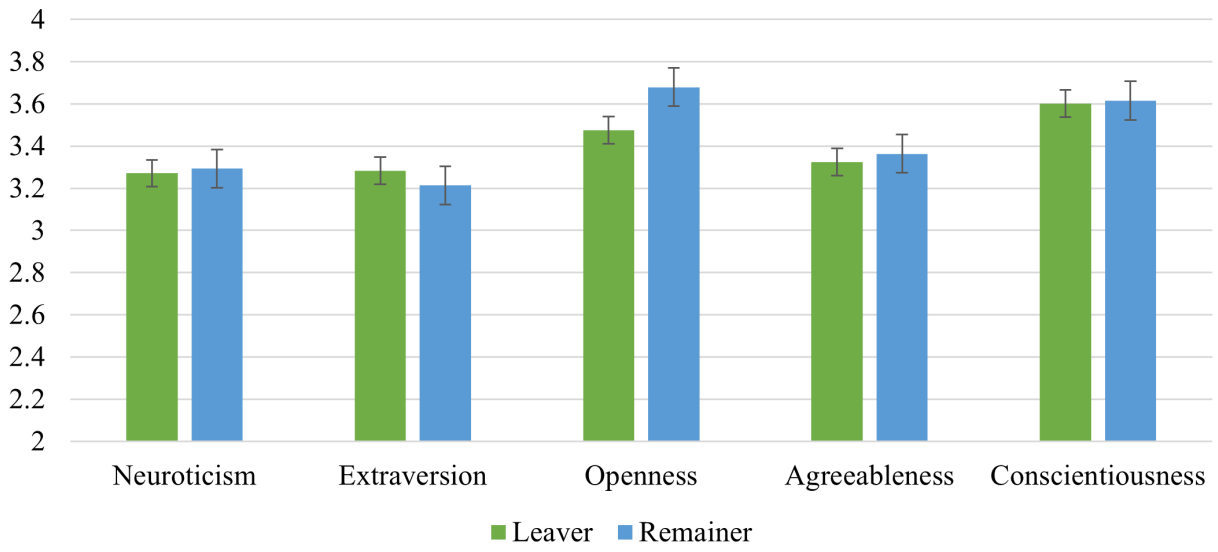


Figure 5.2: Big Five personalities by voting decision

Note: Bars indicate the score of each personality trait. The green bar shows the average Big Five personality traits for Leave voters. The blue bar shows Remain voters' average Big Five personality traits score. Error bars indicate 95% confidence intervals.

In Table 5.2, we present results from an OLS regression model that includes the Big Five measures. We also include a vector of controls. Our dependent variables are mean transfers

pooled across all receivers in the prisoner's dilemma, dictator game, trust game, modified public goods game and public goods game. The coefficients on the interaction term between the Big Five traits and identities (a dummy variable *Leaver*, which equals one if they chose to Remain or 0 if they chose to leave the EU in the 2016 Referendum) identify the extent that Leavers in comparison to Remainers behave depending on the different personality traits. Individuals who report being highly neurotic coordinate more in the prisoner's dilemma (see column 1).

Consistent with our hypothesis, we find that extraverted participants contribute more in the public goods game (see column 4). However, being highly extraverted when the participant is a Leave voter reduces contributions in the public goods game. Therefore, as presented in the public goods game, Extraversion strongly predicts prosocial behaviours in the group setting for Remainers.

A high Openness rating increases contributions in the trust game (see column 3). Leave participants drive this cooperativeness. Thus, Openness is more positively associated with greater trust for Leave voters than Remainers. This contradicts our hypothesis; however, further empirical analysis shows that Leavers' contributions drive this to UK receivers. Therefore, this may be indicative of wanting to solidify a British identity. The results of highly agreeable individuals are consistent with our conceptual framework. Highly Agreeable individuals are found to be more trusting. Lastly, high levels of Conscientiousness did not significantly impact cooperative behaviour [7].

[7] To account for any nonlinear relationship, we replicate estimates in Table 5.2 including the squared terms on personality traits, see Appendix D, Table D1. As a further robustness test, we replicate regression in the main text, including each of the 15-item personality traits, see Appendix D, Table D2. In addition, previous research has argued that political conservatism is associated with personality traits such as high conscientiousness (Hirsh et al. 2010, Sibley et al. 2012, Jost et al. 2003). Therefore, as a further robustness check, we run personality traits and political orientation using the 2019 General Election on mean contributions, see Appendix D, Table D2.

Table 5.2: Regression for personality effects in behaviour

Dep. variable: Aggregated amount sent	PD	DG	TG	PGG	mPGG
	(1)	(2)	(3)	(4)	(5)
Leaver	1.495	-0.155	-0.148	0.549	-1.064
	(1.551)	(1.389)	(1.343)	(1.371)	(1.548)
Neuroticism	0.293*	-0.043	-0.144	-0.055	-0.051
	(0.167)	(0.156)	(0.165)	(0.160)	(0.185)
Neuroticism x Leaver	-0.362	0.029	-0.120	-0.086	0.144
	(0.272)	(0.208)	(0.239)	(0.207)	(0.228)
Extraversion	0.145	-0.188	-0.164	0.284*	0.094
	(0.176)	(0.163)	(0.153)	(0.154)	(0.206)
Extraversion x Leaver	-0.190	0.087	0.024	-0.466**	-0.222
	(0.252)	(0.276)	(0.222)	(0.204)	(0.263)
Openness	-0.107	-0.124	-0.051	-0.055	0.012
	(0.131)	(0.110)	(0.100)	(0.117)	(0.131)
Openness x Leaver	0.158	0.220	0.259*	0.020	0.012
	(0.172)	(0.144)	(0.143)	(0.136)	(0.162)
Agreeableness	-0.027	0.163	0.286*	-0.235	-0.009
	(0.191)	(0.178)	(0.160)	(0.173)	(0.238)
Agreeableness x Leaver	-0.051	-0.209	-0.221	0.183	0.008
	(0.279)	(0.233)	(0.228)	(0.211)	(0.278)
Conscientiousness	0.194	0.002	-0.182	0.001	0.113
	(0.216)	(0.176)	(0.178)	(0.211)	(0.251)
Conscientiousness x Leaver	-0.027	-0.077	0.086	0.153	0.325
	(0.304)	(0.265)	(0.271)	(0.260)	(0.315)
Controls	YES	YES	YES	YES	YES
Constant	1.029	2.248	2.069	1.821	-0.666
	(1.640)	(1.570)	(1.533)	(1.360)	(1.632)
N	504	505	505	505	505

Note: The dependent variable is the pooled mean contributions for all receivers (leave, remain, and UK) in each game. We add demographic controls and beliefs. *** p<0.01, **

p<0.05, * p<0.10.

To clearly disentangle the direction and magnitude of prosocial behaviour on personality traits. We run separate regression models with the dependent variable, the amount sent to each receiver. Table 5.3 presents the Probit model by personality traits in the prisoner's dilemma. The interaction between identity and Neuroticism tells us that highly neurotic Leavers are more likely to defect when the other player is a Remainer. Similarly, highly conscientious Leavers chose to defect with Remainers. In contrast, agreeable Leavers are more likely to coordinate with Remain receivers. Likewise, highly open Leavers are more likely to coordinate with British receivers. In other words, agreeable and open Leavers display greater levels of unconditional cooperation across political lines. Table 5.4 presents an OLS regression model for the dictator game. Surprisingly, we find that when *Leaver* equals zero, therefore, Open Remainers reduces the amount sent to Remainers.

Similarly, in the trust game, we observe that Openness reduces the amount sent to Remainers, see Table 5.5. However, open Leavers cooperate with British receivers[8]. In addition, extraverted Leavers are highly trusting towards fellow Leavers. We also find that more agreeable participants increase trust and amounts sent to Leave receivers.

Table 5.6 presents an OLS analysis in the public goods game. We observe that highly extraverted Remainers contribute more when the other players are Receivers. When *Leaver* equals zero, therefore, when the participant is a Remainer, we observe lower levels of cooperative behaviour with British recipients. In the modified public goods game, highly conscientious Leavers contribute more to the common pot when paired with a Remainer and British player, see Table 5.7.

[8]The unexpected direction of the coefficient is mirrored in other studies which similarly find converse relationships with personality traits and prosocial behaviour. For example, Ben-Ner & Kramer (2011) finds inconsistent results regarding extraversion and prosocial behaviour. See Kline et al. (2019) for multilevel meta-analysis, who observe inconsistencies in the literature, even within the same paper, as in our case, but find a robust result following the meta-analysis of Openness.

Table 5.3: Probit regression for personality effects in behaviour (Prisoner's Dilemma)

Dep. variable: Coordinate(=1, 0 otherwise)	Leaver	Remainer	British
	(1)	(2)	(3)
Leaver	1.169	2.200	-0.071
	(1.380)	(1.425)	(1.501)
Neuroticism	0.061	0.327**	0.055
	(0.146)	(0.150)	(0.147)
Neuroticism x Leaver	-0.130	-0.417**	0.068
	(0.214)	(0.212)	(0.220)
Extraversion	0.113	0.019	0.055
	(0.159)	(0.169)	(0.163)
Extraversion x Leaver	-0.153	-0.034	-0.163
	(0.225)	(0.236)	(0.235)
Openness	-0.034	-0.006	-0.069
	(0.110)	(0.116)	(0.106)
Openness x Leaver	0.104	-0.093	0.259*
	(0.150)	(0.154)	(0.154)
Agreeableness	0.207	-0.214	-0.216
	(0.163)	(0.184)	(0.177)
Agreeableness x Leaver	-0.351	0.455*	0.044
	(0.234)	(0.242)	(0.241)
Conscientiousness	-0.075	0.148	0.128
	(0.175)	(0.188)	(0.180)
Conscientiousness x Leaver	0.232	-0.664**	-0.122
	(0.271)	(0.279)	(0.279)
Controls	YES	YES	YES
Constant	-0.380	-1.354	0.561
	(1.337)	(1.409)	(1.378)
N	504	504	504

Note: The dependent variable is the likelihood of coordination(=1, 0 otherwise) for each receiver. We add demographic controls and beliefs. *** p<0.01, ** p<0.05, * p<0.10.

Table 5.4: OLS regression for personality effects in behaviour (Dictator Game)

Dep. variable: Amount sent to receiver	Leaver	Remainer	British
	(1)	(2)	(3)
Leaver	0.726	0.355	-1.060
	(2.126)	(2.193)	(1.828)
Neuroticism	-0.149	0.320	-0.002
	(0.198)	(0.201)	(0.188)
Neuroticism x Leaver	0.061	-0.300	0.095
	(0.277)	(0.304)	(0.271)
Extraversion	-0.115	-0.166	-0.108
	(0.247)	(0.198)	(0.226)
Extraversion x Leaver	-0.033	0.097	0.020
	(0.386)	(0.313)	(0.364)
Openness	-0.159	-0.240*	-0.179
	(0.157)	(0.132)	(0.139)
Openness x Leaver	0.322	0.275	0.251
	(0.211)	(0.185)	(0.188)
Agreeableness	0.297	0.115	0.176
	(0.221)	(0.225)	(0.224)
Agreeableness x Leaver	-0.243	-0.231	-0.207
	(0.345)	(0.336)	(0.318)
Conscientiousness	-0.015	-0.136	-0.115
	(0.277)	(0.229)	(0.240)
Conscientiousness x Leaver	-0.283	0.030	0.098
	(0.431)	(0.373)	(0.394)
Controls	YES	YES	YES
Constant	3.647*	1.655	2.530
	(2.035)	(1.666)	(1.754)
N	505	505	505

Note: The dependent variable is the amount sent to each receiver. We add demographic controls and beliefs. *** p<0.01, ** p<0.05, * p<0.10.

Table 5.5: OLS regression for personality effects in behaviour (Trust Game)

Dep. variable: Amount sent to receiver	Leaver	Remainer	British
	(1)	(2)	(3)
Leaver	-1.342	1.366	-2.459
	(1.631)	(2.191)	(1.824)
Neuroticism	-0.173	0.009	-0.051
	(0.214)	(0.237)	(0.183)
Neuroticism x Leaver	0.193	-0.300	-0.196
	(0.329)	(0.341)	(0.305)
Extraversion	-0.292	-0.041	-0.055
	(0.191)	(0.237)	(0.170)
Extraversion x Leaver	0.634**	-0.581	0.099
	(0.287)	(0.365)	(0.268)
Openness	0.119	-0.279*	-0.135
	(0.144)	(0.152)	(0.114)
Openness x Leaver	0.054	0.332	0.562***
	(0.219)	(0.223)	(0.191)
Agreeableness	0.429**	0.149	0.142
	(0.207)	(0.230)	(0.183)
Agreeableness x Leaver	-0.450	-0.087	-0.123
	(0.305)	(0.385)	(0.296)
Conscientiousness	-0.379	-0.064	-0.247
	(0.236)	(0.274)	(0.180)
Conscientiousness x Leaver	-0.066	0.181	0.335
	(0.477)	(0.464)	(0.433)
Controls	YES	YES	YES
Constant	1.920	1.963	2.773
	(1.937)	(2.044)	(1.736)
N	505	505	505

Note: The dependent variable is the amount sent to each receiver. We add demographic controls and beliefs. *** p<0.01, ** p<0.05, * p<0.10.

Table 5.6: OLS regression for personality effects in behaviour (Public Goods Game)

Dep. variable: Amount sent to receiver	Leaver	Remainer	British
	(1)	(2)	(3)
Leaver	1.459	-0.209	-0.106
	(2.299)	(2.401)	(1.536)
Neuroticism	-0.202	0.150	0.142
	(0.230)	(0.207)	(0.179)
Neuroticism x Leaver	-0.078	-0.252	-0.396
	(0.343)	(0.349)	(0.274)
Extraversion	0.158	0.354*	0.265
	(0.230)	(0.205)	(0.175)
Extraversion x Leaver	-0.359	-0.528	-0.401
	(0.311)	(0.322)	(0.251)
Openness	0.044	-0.202	-0.134
	(0.170)	(0.139)	(0.129)
Openness x Leaver	0.128	-0.018	0.007
	(0.225)	(0.208)	(0.181)
Agreeableness	-0.140	-0.307	-0.382**
	(0.247)	(0.233)	(0.176)
Agreeableness x Leaver	-0.260	0.399	0.092
	(0.333)	(0.371)	(0.248)
Conscientiousness	-0.107	0.005	-0.120
	(0.264)	(0.255)	(0.219)
Conscientiousness x Leaver	0.133	0.397	0.636
	(0.522)	(0.530)	(0.437)
Controls	YES	YES	YES
Constant	4.000*	1.749	1.742
	(2.155)	(1.962)	(1.579)
N	505	505	505

Note: The dependent variable is the amount sent to each receiver. We add demographic controls and beliefs. *** p<0.01, ** p<0.05, * p<0.10.

Table 5.7: OLS regression for personality effects in behaviour (Modified Public Goods Game)

Dep. variable: Amount sent to receiver	Leaver	Remainer	British
	(1)	(2)	(3)
Leaver	-1.527	0.034	-2.628
	(2.320)	(1.760)	(1.969)
Neuroticism	-0.091	0.170	-0.005
	(0.254)	(0.208)	(0.221)
Neuroticism x Leaver	0.317	-0.094	-0.074
	(0.322)	(0.281)	(0.276)
Extraversion	0.022	0.135	0.219
	(0.280)	(0.218)	(0.226)
Extraversion x Leaver	-0.002	-0.475	-0.355
	(0.346)	(0.315)	(0.326)
Openness	-0.131	-0.073	0.073
	(0.203)	(0.146)	(0.161)
Openness x Leaver	0.287	-0.043	-0.137
	(0.246)	(0.205)	(0.203)
Agreeableness	0.306	0.022	-0.371
	(0.327)	(0.234)	(0.259)
Agreeableness x Leaver	-0.308	-0.227	0.566
	(0.387)	(0.334)	(0.349)
Conscientiousness	0.259	-0.244	-0.205
	(0.331)	(0.238)	(0.235)
Conscientiousness x Leaver	0.164	0.735*	0.750*
	(0.412)	(0.410)	(0.383)
Controls	YES	YES	YES
Controls	-0.039	1.327	1.788
	(2.359)	(1.622)	(1.913)
N	505	505	505

Note: The dependent variable is the amount sent to each receiver. We add demographic controls and beliefs. *** p<0.01, ** p<0.05, * p<0.10.

5.4.2 Misrepresenting voting decision

We now turn to the underlying personality types related to participants' inconsistent identification with the experimenter and online recruitment platforms. In Table 5.8, we run a Probit regression model on the Big Five personality traits. Our dependent variable is a dummy variable equal to one if participants report the same voting decision between the experimenter and online platform and zero otherwise. Extraverted Remainers are more likely to have reported the same identity prior to and during the experiment. Conscientious participants are likely to consistently report their identity to the experimenter and the platform, whereas highly Conscientious Leavers are more likely to misrepresent. In other words, highly Conscientious Leavers report in our experiment that they were a Leaver when they informed Prolific that they voted to Remain. Our results may indicate that Remainers (as identified to Prolific) may have feelings of regret and embarrassment.

Table 5.8: Probit regression misrepresenting voting decision

Dep. variable: correct identity (=1, 0 otherwise)	(1)	(2)
Leaver	1.154	1.129
	(1.277)	(1.378)
Neuroticism	-0.102	-0.161
	(0.136)	(0.150)
Neuroticism x Leaver	0.104	0.073
	(0.196)	(0.207)
Extraversion	0.281*	0.305**
	(0.146)	(0.155)
Extraversion x Leaver	-0.041	-0.057
	(0.210)	(0.223)
Openness	-0.090	-0.108
	(0.103)	(0.110)
Openness x Leaver	0.048	0.092
	(0.140)	(0.150)
Agreeableness	-0.086	-0.111
	(0.161)	(0.172)
Agreeableness x Leaver	0.191	0.122
	(0.220)	(0.238)
Conscientiousness	0.335**	0.346*
	(0.165)	(0.185)
Conscientiousness x Leaver	-0.623**	-0.547**
	(0.249)	(0.269)
Controls	NO	YES
Constant	-1.300	-4.504***
	(0.985)	(1.383)
N	505	504

Note: The dependent variable is a dummy variable that equals 1 if participants report the same identity to the experimenter and prescreening criteria 0 otherwise. In column (2), we add demographic controls. *** p<0.01, ** p<0.05, * p<0.10.

5.5 Discussion and conclusion

The Big Five dimensions are a highly-validated instrument for measuring personality traits. Previous research finds that different personality traits are relevant to political orientation and decision-making (Hirsh et al. 2010, Jost et al. 2003, Rentfrow et al. 2009). We contribute to this literature by identifying the associated personality traits of British citizens and the subsequent influence of cooperative behaviour and identity reporting in experiments. The outcome of the EU Referendum highlighted apparent differences between citizens regarding integration, globalisation, and immigration. Understanding the personality traits of Leave and Remain voters can assist policymakers in understanding the underlying predictors that lead individuals to support specific political outcomes. Moreover, given the context of political polarisation (Hanel & Wolf 2020, McCoy et al. 2018), understanding the personality traits of voters' decisions can shed light on how these groups will engage and interact with one another and other politically divisive issues.

We set out to gain an understanding of the individual inter-personality differences between Leavers and Remainers on cooperative behaviour. Similar to studies conducted in America (see Hirsh et al. (2010) for example), we find differences in individual-level characteristics between opposing political identities. Specifically, we observe that Leavers and Remainers differ in terms of Openness. Our finding supports the influence of Euroscepticism (Hobolt et al. 2021), where Leavers fear the consequences of integration with other cultures, whereas, Remainers are more open to change and the spillovers of globalisation. We also hypothesised that individual variation in personality traits would affect in-group favouritism and out-group discrimination (Swope et al. 2008, Ben-Ner & Kramer 2011). We find some evidence that personality traits of Leavers and Remainers influence behavioural decisions when engaging with Leavers, Remainers, and British persons. Our findings suggest that Openness can partly explain political orientation and play a role in cooperative behaviour, as displayed in a dictator game and trust game. Findings from the prisoner's dilemma are consistent with (Hirsh & Peterson 2009), that neuroticism influences coordination in our case, increasing coordination with in-group receivers and defection with out-group receivers. These findings are consistent with the behavioural economic research that find that individuals fear punishment (Fehr & Gachter 2000, Abbink et al. 2010, Henrich et al. 2016).

In addition, Gerber et al. (2013) find that less agreeable individuals are more likely to be dishonest about their voter turnout. Whether deliberately or not, we assume that personality traits may play a role in an individual misreporting the decision to Leave or Remain.

Even after controlling for demographic factors, we find that Extraversion and Conscientiousness are significant predictors of misrepresenting their identity between the online platform and the experimenter. However, the results may be due to participants' errors regarding misreporting. Conscientious Leavers are more likely to have misstated Remain to Prolific but Leave in the experiments. Therefore, consistent with Hart et al. (2020), we find that being an extraverted Remainer is associated with less dishonesty, in other words, with less chance of misrepresenting their voting decision. Although our results need to be read with caution, we speculate that the finding suggests that Conscientious voters who chose to Remain (as reported to the recruitment platform) subsequently may regret their decision to vote remain, perhaps due to facets of conscientiousness being associated with willingness to achieve making them regretful (Costa et al. 1991) of being on the losing side and therefore choose to identify as a Leaver. Our finding complements the literature on conscientiousness and 'authentic pride'. Our findings suggest that either through pride or regret, Conscientious participants misstate their identity (Carver et al. 2010, Cheng et al. 2010).

Our findings support the growing literature on misrepresentation in online experiments. In our experiment, there is no benefit to misstating one's identity, as prescreening has occurred prior to the experiment. The online inclusion criteria consisted of whether the participant had voted rather than any emphasis on what they voted. Therefore, lying about 'what' you voted for is no benefit to gaining access to the experiment. Given that simply voting was sufficient, the relationship with personalities and what participants reported to the platform and during the experiment provides some means of understanding misrepresentation behaviours. Given the choice environment and context of our research, participants may misreport due to experimenter effect or because they deem one identity as more "personally or socially desirable" (Chandler & Paolacci 2017, p.505).

The paper does, however, suffer from some limitations. The major limitation is that we have a low-reliability alpha score when measuring personalities. Although each of the 15-item markers yields a satisfactory Cronbach's alpha, once grouped into the five dimensions, our measures suffer from lower reliability. Therefore, we assume some personality traits that influence behaviours may be underestimated. However, this partly may be due to the hypothetical choice environment in the games (Thielmann et al. 2020, p.45). Thus, future research can improve on this paper by using a greater sample of participants, conducting laboratory experiments or alternatively using real-world scenarios such as giving to charity. We also acknowledge the limitations of using subjective measures in experimental research (Jahedi & Méndez 2014). In addition, there is always the danger that experiments conducted online are

less reliable, as shown by the low Cronbach alpha and misrepresenting findings. Nonetheless, the results provide an avenue for future research to further understand personality traits in identity research and the potential for inconsistencies in experimental research. Moreover, given that the Big Five are seen as consistent across cultures, future research could extend upon the existing commonly used Western student population and use developing countries (Paunonen & Ashton 2001).

In conclusion, we set out to investigate the association between the Big Five dimensions and behaviours between Leave and Remain voters in the UK. We find that Leavers report lower levels of Openness, which is consistent with the micro-level data, which found that Leavers were less open to immigration and globalisation (Hanel & Wolf 2020, Hobolt 2016, Ashcroft 2016). That said, the result does not mean that all Leavers are low in Openness and therefore close-minded and have limited creativity. Instead, to an extent, Openness can manifest itself as an element driving the decision to have voted to leave the EU. Using the short Big 5, we also observe that personalities explain cooperative behaviour to some degree. We find evidence that personality traits, when interacting with identities, impact favouritism and discriminatory behaviour. In addition, we find evidence that Conscientiousness and Extraversion are predictors of misrepresenting their voting decision during the Referendum. Therefore, researchers must be aware of the potential for low-quality data derived from online experiments.

6

Conclusion

The thesis consists of four chapters that use experimental methods to investigate social identities, the interactions between groups and mechanisms to shift behaviours. In two studies (chapter 2 and chapter 3), we investigated discriminatory behaviour, polarisation and the impacts of norm-based nudges. Further expanding on social preferences, two other chapters (chapter 4 and chapter 5) considered in which environments individuals will lie for others and issues related to personality traits behind cooperative behaviour, respectively. We also considered the issue experimenters face when participants "choose" an identity.

In chapter 2, we examined inter-group preferences in Zimbabwe between the two main ethnolinguistic groups - Ndebele and Shona. We considered five main conjectures: (i) identity-dependent behaviours, (ii) beliefs about others' behaviours, (iii) salience of group identities or polarisation, (iv) social norm conformity and (v) can these preferences be shifted. Various experimental games were utilised in light of the existing literature, with crucial modifications contributing to the literature. We used the social norms elicitation task adapted from Krupka & Weber (2013) to identify differences between personal and group norms (Mackie 1986), as well as norm effects on discriminatory behaviour. In chapter 3, we also investigated

social identities and discriminatory behaviour, using a similar set of social dilemma games, salience questions, social norms and norm-based nudges. In comparison to chapter 2, we looked at non-partisan opinion-based identities - Leavers and Remainers in the United Kingdom. Consistent with previous research (see, for example, McCoy et al. (2018) and Iyengar & Westwood (2015)), we found evidence of discriminatory behaviour between Leavers and Remainers. The research has also shown that the non-rival group in the study conducted in the United Kingdom did not result in feelings of hostility or favouritism (Abbink & Harris 2019). Therefore, policy-makers who are interested in minimising discrimination can emphasise the "UK" identity.

If these social identities are significant enough, they can begin to govern how people ought to behave (Chen et al. 2014, Hobolt et al. 2021). We argued that as groups become more polarised, there will be greater affiliation with the in-group. We found that group salience, via channels such as self-identification or feeling highly aligned to the identity, significantly influenced the tendency for cooperative behaviour (Eckel & Grossman 2005, Chen & Chen 2011, Böhm et al. 2020). Our findings are consistent with the literature that finds that social identity salience is relevant in discriminatory behaviour (Abbink & Harris 2012, Ahmed 2010, Fong & Luttmer 2011, Chmura et al. 2016). Furthermore, in both settings, we found evidence of pluralistic ignorance. The dangers of pluralistic ignorance have been previously documented in the literature, which can further exacerbate divisions (Smerdon et al. 2020, Miller & Prentice 2016, Bursztyn et al. 2020).

In these two chapters, we also used social norm-based interventions to examine if the gap between the polarised groups can be further divided or healed. We combine prescriptive and proscriptive norms to distinguish what information matters in social norm-based nudges. In Zimbabwe, we used norms around interethnic marriages. In the United Kingdom, we used norms are vandalism. We compared our treatments with no social norm-based intervention. In general, we found that both norm nudges altered behaviours, although the magnitude and direction of the effect was not always consistent. In these chapters, we contribute to the experimental literature by identifying a population gap by conducting experiments in a developing country like Zimbabwe. The research also provides further insights into the influence of norm-based interventions on behaviours across different contexts.

A limitation of these two studies is that our social norm-based interventions fail to deliver a clear pattern of behaviour. However, our finding is consistent with the existing literature that has also observed mixed results of norm nudges (see, for example, Dimant et al. (2020)). Therefore, further research can consider how interventions such as ours can be used in policy,

perhaps as a part of a package alongside other tools. Furthermore, our findings regarding polarisation and strength of group affiliation provide an avenue for future research to consider other social salience questionnaires.

In chapter 4, the propensity for honesty for oneself and others was tested. In this chapter, we show that group identity and the salience of the group identity impact dishonesty by increasing the moral cost of dishonest behaviour. A secondary finding from this chapter is that social norms around bribery shape dishonest behaviour. An additional noteworthy result is the importance of trust in institutions. The extent that participants trusted the government significantly impacted dishonesty on behalf of others in both the United Kingdom and Zimbabwe. Therefore, an extension of the research could further investigate the role of institutions and government scandals on citizens' general acceptance of dishonest behaviours. Further implications are discussed in the chapter.

Lastly, this thesis also attempts to fill a method gap in considering the issue of reporting identities in in-group and out-group experiments. We correlate this with personalities. Chapter 5 examined personality traits and prosocial behaviours. We set out to investigate the association between the Big Five taxonomies and behaviours between Leave and Remain voters in the UK. We find that Leavers report lower levels of Openness, which is consistent with established micro-level data (see, for example, Alabrese et al. (2019) and Hobolt (2016)), which found that Leavers were less open to immigration and globalisation. (Hanel & Wolf 2020, Hobolt 2016, Ashcroft 2016). This finding suggests that Openness characteristics are linked to Euroscepticism. We also observed that personalities do explain, to some degree, cooperative behaviour. An interesting finding from this chapter is that Conscientiousness and Extraversion are predictors of misrepresenting their social categorisation. Given the identity being drawn upon, this result may be evidence that highly Conscientious Remainers (reported to the recruitment service, but Leavers in our experiment) may regret their decision. The findings from this chapter contribute to the experimental, political and psychological literature. Further research can pursue these themes using a greater sample of participants, other recruitment platforms, and other salient identities.

In conclusion, like Zimbabwe, Britain has emphasised in and out-groups in political endeavours. Indeed, the manipulation of in and out-groups in the political domain is not unique to Zimbabwe or the UK (Ndhlovu 2007, Mounk 2018, McCoy et al. 2018). Nonetheless, these two nations exhibit similar patterns of polarisation. This study set out to offer some critical insight into these two contexts, the persistence of polarisation and the extent of the group affiliation (Bernhard et al. 2006). Throughout the thesis, self-identification and the strength

of the association with the group is paramount, in other words, how participants feel about their identity. Our findings have policy implications in a wide variety of contexts. As countries face greater polarisation and these divisions spill into the public sphere. If unchecked, this could prove a real challenge and may hinder rather than facilitate democratic practices.

Bibliography

Abbink, K., Brandts, J., Herrmann, B. & Orzen, H. (2010), 'Intergroup confict and intra-group punishment in an experimental contest game', *American Economic Review* **100**(1), 420–447.

Abbink, K. & Harris, D. (2012), 'In-Group Favouritism and Out-Group Discrimination in Naturally Occurring Groups', *University of Oxford: Discussion Paper Series* (1471-0498), 1–17.

Abbink, K. & Harris, D. (2019), 'In-group favouritism and out-group discrimination in naturally occurring groups', *PLOS ONE* **14**(9), 1–13.
URL: *https://doi.org/10.1371/journal.pone.0221616*

Abbink, K. & Hennig-Schmidt, H. (2006), 'Neutral versus loaded instructions in a bribery experiment', *Experimental Economics* **9**(2), 103–121.

Abeler, J., Becker, A. & Falk, A. (2014), 'Representative evidence on lying costs', *Journal of Public Economics* **113**, 96–104.
URL: *http://dx.doi.org/10.1016/j.jpubeco.2014.01.005*

Abeler, J., Nosenzo, D. & Raymond, C. (2019), 'Preferences for Truth-Telling', *Econometrica* **87**(4).

Abrahamse, W., Steg, L., Vlek, C. & Rothengatter, T. (2005), 'A review of intervention studies aimed at household energy conservation', *Journal of Environmental Psychology* **25**(3), 273–291.

Acemoglu, D. & Jackson, M. O. (2017), 'Social norms and the enforcement of laws', *Journal of the European Economic Association* **15**(2), 245–295.

Achilles, S. L., Mhlanga, F. G., Musara, P., Poloyac, S. M., Chirenje, Z. M. & Hillier, S. L. (2018), 'Misreporting of contraceptive hormone use in clinical research participants', *Contraception* **97**(4), 346–353.
URL: *https://doi.org/10.1016/j.contraception.2017.09.013*

Ahmed, A. M. (2010), 'What is in a surname? The role of ethnicity in economic decision making', *Applied Economics* **42**(21), 2715–2723.

Ahmed, A. M. & Salas, O. (2009), 'Is the hand of God involved in human cooperation?', *International Journal of Social Economics* **36**(1-2), 70–80.

Akerlof, G. A. & Kranton, R. E. (2000), 'Economics and Identity', *The Quarterly Journal of Economics* **115**(3), 715–753.

Akerlof, G. A. & Kranton, R. E. (2005), 'Identity and the Economics of Organizations', *Journal of Economic Perspectives* **19**(1), 9–32.

Alabrese, E., Becker, S. O., Fetzer, T. & Novy, D. (2019), 'Who voted for Brexit? Individual and regional data combined', *European Journal of Political Economy* **56**(August 2018), 132–150.
URL: *https://doi.org/10.1016/j.ejpoleco.2018.08.002*

Albornoz, F., Bradley, J. & Sonderegger, S. (2021), Brexit and hate crime: why was the rise more pronounced in areas that voted Remain?, Technical report, LSE British Politics and Policy.
URL: *https://blogs.lse.ac.uk/politicsandpolicy/brexit-and-hate-crime/*

Alesina, A., Devleeschauwer, A., Easterly, W. & Kurlat, S. (8), 'Fractionalization', *Journal of Economic Growth* pp. 155–194.

Alford, J. R. & Hibbing, J. R. (2007), 'Personal, interpersonal, and political temperaments', *Annals of the American Academy of Political and Social Science* **614**(1), 196–212.

Ali, S. N. & Bénabou, R. (2016), 'Image Versus Information', *NBER Working Paper* **22203**.

Allcott, H. (2011), 'Social norms and energy conservation', *Journal of Public Economics* **95**(9-10), 1082–1095.
URL: *http://dx.doi.org/10.1016/j.jpubeco.2011.03.003*

Anderson, L. R. & Mellor, J. M. (2009), 'Religion and cooperation in a public goods experiment', *Economics Letters* **105**(1), 58–60.
URL: *http://dx.doi.org/10.1016/j.econlet.2009.05.016*

Andor, M. A., Gerster, A. & Peters, J. (2022), 'Information campaigns for residential energy conservation', *European Economic Review* p. 104094.
URL: *https://doi.org/10.1016/j.euroecorev.2022.104094*

Apffelstaedt, A., Freundt, J. & Oslislo, C. (2022), 'Social norms and elections: How elected rules can make behavior (in)appropriate', *Journal of Economic Behavior & Organization* **196**, 148–177.

Arad, A. & Rubinstein, A. (2018), 'The People ' s Perspective on Libertarian-Paternalistic Policies', *Journal of Law and Economics* **61**(2), 311–333.

Arbel, Y., Bar-El, R., Siniver, E. & Tobol, Y. (2014), 'Roll a die and tell a lie - What affects honesty?', *Journal of Economic Behavior and Organization* **107**(PA), 153–172.
URL: *http://dx.doi.org/10.1016/j.jebo.2014.08.009*

Arrow, K. J. (1972), Some mathematical models of race discrimination in the labor market, *in* A. Pascal, ed., 'Racial Discrimination in Economic Life', D.C. Heath, Lexington, MA:, pp. 187–204.

Ashcroft, M. (2016), 'How the United Kingdom Voted on Thursday...and Why'.
URL: *https://lordashcroftpolls.com/2016/06/how-the-united-kingdom-voted-and-why/*

Azmanova, A. (2011), 'After the Left-Right (Dis)continuum: Globalization and the Remaking of Europe's Ideological Geography', *International Political Sociology* **5**(4), 384–407.

Baliga, S., Hanany, E. & Klibanoff, P. (2013), 'Polarization and Ambiguity', *American Economic Association* **103**(7), 3071–3083.

Balliet, D., Wu, J. & De Dreu, C. K. (2014), 'Ingroup favoritism in cooperation: a meta-analysis', *Psychological bulletin* **140**(6), 1556–1581.

Barak-Corren, N. (2021), 'Regulating for integration by behavioral design: An evidence-based approach for culturally responsive regulation', *Regulation and Governance* **15**(S1), S163–S182.

Barber, T. X. & Silver, M. J. (1968), 'Fact, fiction, and the experimenter bias effect', *Psychological Bulletin* **70**(6 Part 2), 1–29.

Barr, A., Lane, T. & Nosenzo, D. (2018), 'On the social inappropriateness of discrimination', *Journal of Public Economics* **164**, 153–164.
URL: *https://doi.org/10.1016/j.jpubeco.2018.06.004*

Barthelemy, J. J. & Lounsbury, J. W. (2009), 'The relationship between aggression and the big five personality factors in predicting academic success', *Journal of Human Behavior in the Social Environment* **19**(2), 159–170.

Baumeister, R. F., Vohs, K. D. & Funder, D. C. (2007), 'Psychology as the Science of Self-Reports and Finger Movements: Whatever Happened to Actual Behavior?', *Perspectives on Psychological Science* **2**(4), 396–403.

Becker, G. S. (1968), 'Crime and Punishment : An Economic Approach', *Journal of Political Economy* **76**(2), 169–217.

Becker, G. S. (1971), *The Economics of Discrimination*, 2nd edn, Chicago Press, Chicago.

Belot, M., Duch, R. & Miller, L. (2015), 'A comprehensive comparison of students and non-students in classic experimental games', *Journal of Economic Behavior and Organization* **113**, 26–33.

Ben-Ner, A. & Halldorsson, F. (2010), 'Trusting and trustworthiness: What are they, how to measure them, and what affects them', *Journal of Economic Psychology* **31**(1), 64–79.
URL: *http://dx.doi.org/10.1016/j.joep.2009.10.001*

Ben-Ner, A., Kong, F. & Putterman, L. (2004), 'Share and share alike? Gender-pairing, personality, and cognitive ability as determinants of giving', *Journal of Economic Psychology* **25**(5), 581–589.

Ben-Ner, A. & Kramer, A. (2011), 'Personality and altruism in the dictator game: Relationship to giving to kin, collaborators, competitors, and neutrals', *Personality and Individual Differences* **51**(3), 216–221.
URL: *http://dx.doi.org/10.1016/j.paid.2010.04.024*

Ben-Ner, A., McCall, B. P., Stephane, M. & Wang, H. (2009), 'Identity and in-group/out-group differentiation in work and giving behaviors: Experimental evidence', *Journal of Economic Behavior and Organization* **72**(1), 153–170.

Bénabou, R. & Tirole, J. (2011*a*), 'Identity, morals, and taboos: Beliefs as assets', *The Quarterly Journal of Economics* **126**(2), 805–855.

Bénabou, R. & Tirole, J. (2011*b*), 'Laws and Norms', *NBER Working Paper* .

Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., Tucker-Ray, W., Congdon, W. J. & Galing, S. (2017), 'Should Governments Invest More in Nudging?', *Psychological Science* **28**(8), 1041–1055.

Benistant, J. & Villeval, M. C. (2019), 'Unethical behavior and group identity in contests', *Journal of Economic Psychology* **72**, 128–155.
URL: *https://doi.org/10.1016/j.joep.2019.03.001*

Benjamin, D. J., Choi, J. J. & Fisher, G. (2016), 'Religious Identity and Economic Behavior', *The Review of Economics and Statistics* **98**(4), 254 – 267.

Bentham, J. (1781), *An introduction to the Principles of Morals and Legislation*, Clarence Press, Oxford.

Berg, J., Dickhaut, J. & McCabe, K. (1995), 'Trust, Reciprocity, and Social History', *Games and Economic Behavior* **10**(1), 122–142.

Bernhard, H., Fehr, E. & Fischbacher, U. (2006), 'Group Affiliation and Altruistic Norm Enforcement', *American Economic Review* **96**(2), 217–221.

Bicchieri, C. (2015), Why Do People Do What They Do?: A Social Norms Manual for Zimbabwe and Swaziland, Technical report, Innocenti Toolkit Guide from the UNICEF Office of Research, Florence, Italy.
URL: *https://www.unicef-irc.org/research/pdf/429-Social-Norms-Manual-ESARO.pdf*

Bicchieri, C. & Dimant, E. (2019), 'Nudging with care: the risks and benefits of social information', *Public Choice* pp. 1–22.
URL: *https://doi.org/10.1007/s11127-019-00684-6*

Bicchieri, C., Dimant, E., Gächter, S. & Nosenzo, D. (2022), 'Social proximity and the erosion of norm compliance', *Games and Economic Behavior* **132**, 59–72.
URL: *https://doi.org/10.1016/j.geb.2021.11.012*

Bicchieri, C., Dimant, E. & Sonderegger, S. (2021), 'It's Not a Lie If You Believe the Norm Does Not Apply: Conditional Norm-Following with Strategic Beliefs', *SSRN Electronic Journal* .

Bidner, C. & Francois, P. (2011), 'Cultivating Trust: Norms, Institutions and the Implications of Scale', *Economic Journal* **121**(555), 1097–1129.

Bigoni, M., Bortolotti, S., Casari, M. & Gambetta, D. (2019), 'At the root of the north–south cooperation gap in Italy: Preferences or beliefs?', *Economic Journal* **129**(619), 1139–1152.

Birkelund, J., Cherry, T. L. & Mcevoy, D. M. (2022), 'A culture of cheating : The role of worldviews in preferences for honesty', *Journal of Behavioral and Experimental Economics* **96**, 101812.

Bogdanor, V. (2017), 'Would tougher voting rules have changed the Brexit result?'.
URL: *https://www.theguardian.com/commentisfree/2017/oct/10/brexit-eu-referendum-alan-duncan-supermajority-requirement-vote*

Böhm, R., Rusch, H. & Baron, J. (2020), 'The psychology of intergroup conflict: A review of theories and measures', *Journal of Economic Behavior and Organization* **178**, 947–962.
URL: *https://doi.org/10.1016/j.jebo.2018.01.020*

Boles, T. L., Croson, R. T. & Murnighan, J. K. (2000), 'Deception and Retribution in Repeated Ultimatum Bargaining', *Organizational Behavior and Human Decision Processes* **83**(2), 235–259.

Bolton, G., Dimant, E., Schmidt, U. & Robey, S. (2019), 'When a Nudge Backfires : Using Observation with Social and Economic Incentives to Promote Pro-Social Behavior', *CEDEX (No. 2019-03)* .

Bonomi, G., Gennaoili, N. & Tabellini, G. (2021), 'Identity, Beliefs, And Political Conflict', *The Quarterly Journal of Economics* pp. 2371–2411.

Boomgaarden, H. G., Schuck, A. R., Elenbaas, M. & de Vreese, C. H. (2011), 'Mapping EU attitudes: Conceptual and empirical dimensions of Euroscepticism and EU support', *European Union Politics* **12**(2), 241–266.

Boone, C., De Brabander, B. & Van Witteloostuijn, A. (1999), 'The impact of personality on behavior in five Prisoner's Dilemma games', *Journal of Economic Psychology* **20**(3), 343–377.

Börzel, T. A. & Risse, T. (2016), 'Dysfunctional state institutions, trust, and governance in areas of limited statehood', *Regulation and Governance* **10**(2), 149–160.

Bouckaert, J. & Dhaene, G. (2004), 'Inter-ethnic trust and reciprocity: Results of an experiment with small businessmen', *European Journal of Political Economy* **20**(4), 869–886.

Bowles, S. & Polania-Reyes, S. (2012), 'Economic incentives and social preferences: Substitutes or complements?', *Journal of Economic Literature* **50**(2), 368–425.
URL: *https://www.aeaweb.org/articles?id=10.1257/jel.50.2.368*

Brandstätter, H. & Güth, W. (2002), 'Personality in dictator and ultimatum games', *Central European Journal of Operations Research* **10**(3), 191–21.
URL: *http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=7909205&site=ehost-live*

Bruns, H., Kantorowicz-Reznichenko, E., Klement, K., Luistro Jonsson, M. & Rahali, B. (2018), 'Can nudges be transparent and yet effective?', *Journal of Economic Psychology* **65**(March), 41–59.

Bucciol, A. & Piovesan, M. (2011), 'Luck or cheating? A field experiment on honesty with children', *Journal of Economic Psychology* **32**(1), 73–78.

Burger, J. M. (2015), *Personality*, 9 edn, Cengage Learning, Stamford, CT.

Burks, S. V. & Krupka, E. L. (2012), 'A multimethod approach to identifying norms and normative expectations within a corporate hierarchy: Evidence from the financial services industry', *Management Science* **58**(1), 203–217.

Burns, J. (2004), 'Race and trust in post-apartheid South Africa', *Centre for Social Science Research No.78* pp. 1–33.

Bursztyn, L., González, A. L. & Yanagizawa-Drott, D. (2020), 'Misperceived social norms: women working outside the home in saudi arabia', *American Economic Review* **110**(10), 2997–3029.

Bursztyn, L. & Jensen, R. (2017), 'Social Image and Economic Behavior in the Field: Identifying, Understanding, and Shaping Social Pressure', *Annual Review of Economics* **9**(1), 131–153.

Cabinet Office (2022), Investigation into alleged gatherings on government premises during covid restrictions – update 31, Technical Report January, Cabinet Office.

Cadsby, C. B., Du, N. & Song, F. (2016), 'In-group favoritism and moral decision-making', *Journal of Economic Behavior & Organization* **128**, 59–71.
URL: *http://dx.doi.org/10.1016/j.jebo.2016.05.008*

Camerer, C. F. (2011), Introduction, *in* 'Behavioral Game Theory: Experiments in Strategic Interaction', Princeton University Press, Princeton: New Jersey, chapter 1, pp. 1–50.

Caprara, G. V. & Schwartz, S. (2006), 'Personality and Politics : Values , Traits , and', *Political Psychology* **27**(1), 1–28.

Cardenas, J. C. & Carpenter, J. (2008), 'Behavioural development economics: Lessons from field labs in the developing world', *Journal of Development Studies* **44**(3), 311–338.

Cardenas, J. C., Rodriguez, L. A. & Johnson, N. (2010), 'Collective action for watershed management: field experiments in Colombia and Kenya', *Environment and Development Economics* **16**(3), 275–303.

Carlin, R. E. & Love, G. J. (2013), 'The Politics of Interpersonal Trust and Reciprocity: An Experimental Approach', *Political Behavior* **35**(1), 43–63.

Carpenter, J. & Cardenas, J. C. (2011), 'An intercultural examination of cooperation in the commons', *Journal of Conflict Resolution* **55**(4), 632–651.

Carver, C. S., Sinclair, S. & Johnson, S. L. (2010), 'Authentic and hubristic pride: Differential relations to aspects of goal regulation, affect, and self-control', *Journal of Research in Personality* **44**(6), 698–703.
URL: *http://dx.doi.org/10.1016/j.jrp.2010.09.004*

Cason, T. & Mui, V.-L. (1997), 'A Laboratory study of group polarisation in the team Dictator Game', *The Economic Journal* **107**, 1465–1483.

CCJP (1997), The Catholic Commission for Justice and Peace in Zimbabwe & The Legal Resources Foundation, Technical report, The Catholic Commission for Justice and Peace in ZImbabwe & The Legal Resources Foundation.

Chandler, J. J. & Paolacci, G. (2017), 'Lie for a Dime: When Most Prescreening Responses Are Honest but Most Study Participants Are Impostors', *Social Psychological and Personality Science* **8**(5), 500–508.

Charness, G. & Gneezy, U. (2012), 'Strong Evidence for Gender Differences in Risk Taking', *Journal of Economic Behavior and Organization* **83**(1), 50–58.
URL: *http://dx.doi.org/10.1016/j.jebo.2011.06.007*

Chen, B. Y. & Li, S. X. (2009), 'Group Identity and Social Preferences', *American Economic Association* **99**(1), 431–457.

Chen, R. & Chen, Y. (2011), 'The potential of social identity for equilibrium selection', *American Economic Review* **101**(6), 2562–2589.

Chen, Y., Li, S. X., Liu, T. X. & Shih, M. (2014), 'Which hat to wear? Impact of natural identities on coordination and cooperation', *Games and Economic Behavior* **84**, 58–86.
URL: *http://dx.doi.org/10.1016/j.geb.2013.12.002*

Cheng, J. T., Tracy, J. L. & Henrich, J. (2010), 'Pride, personality, and the evolutionary foundations of human social status', *Evolution and Human Behavior* **31**(5), 334–347.
URL: *http://dx.doi.org/10.1016/j.evolhumbehav.2010.02.004*

Chmielewski, M. & Kucker, S. C. (2020), 'An MTurk Crisis? Shifts in Data Quality and the Impact on Study Results', *Social Psychological and Personality Science* **11**(4), 464–473.

Chmura, T., Goerg, S. J. & Weiss, P. (2016), 'Natural groups and economic characteristics as driving forces of wage discrimination', *European Economic Review* **90**, 178–200.
URL: *http://dx.doi.org/10.1016/j.euroecorev.2016.04.005*

Chmura, T., Kube, S., Pitz, T. & Puppe, C. (2005), 'Testing (beliefs about) social preferences : Evidence from an experimental coordination game', *Economic Letters* **88**(2), 214–220.

Chowdhury, S. M., Jeon, J. Y. & Ramalingam, A. (2016), 'Identity and group conflict', *European Economic Review* **90**, 107–121.

Chuah, S. H., Fahoum, R. & Hoffmann, R. (2013), 'Fractionalization and trust in India: A field-experiment', *Economics Letters* **119**(2), 191–194.
URL: *http://dx.doi.org/10.1016/j.econlet.2013.02.029*

Chuah, S.-H., Gächter, S., Hoffmann, R. & Tan, J. H. W. (2023), 'Who discriminates? Evidence from a trust game experiment across three societies', *Journal of Economic Psychology* **97**, 102630.
URL: *https://www.sciencedirect.com/science/article/pii/S0167487023000314*

Chuah, S. H., Hoffmann, R., Jones, M. & Williams, G. (2007), 'Do cultures clash? Evidence from cross-national ultimatum game experiments', *Journal of Economic Behavior and Organization* **64**(1), 35–48.

Cialdini, R. B., Demaine, L. J., Sagarin, B. J., Barrett, D. W., Rhoads, K. & Winter, P. L. (2006), 'Managing social norms for persuasive impact', *Social Influence* **1**(1), 3–15.

Cialdini, R. B., Kallgren, C. A. & Reno, R. R. (1991), 'A focus theory of normative conduct. A theoretical refinement and reevaluation of the role of norms in human behavior', *Advances in experimental social psychology* **24**, 201–234.

Ciriolo, E., Sousa Loureno & Rafael Almeida, S. (2019), The application of behavioural insights to policy in Europe, *in* H. StraSSheim & S. Beck, eds, 'Handbook of Behavioural

Change and Public Policy', Edward Elgar, Cheltenham, UK; Massachusetts, USA, chapter The applic, pp. 102–121.

Cohn, A., Maréchal, M. A., Tannenbaum, D. & Zünd, C. L. (2019), 'Civic honesty around the globe', *Science* **365**(6448), 70–73.

Comparative Constitution Project (2018), 'Zimbabwe's Constitution of 2013'.
URL: *https://www.constituteproject.org/constitution/Zimbabwe_2013.pdf*

Congress of the United States of America (2019), 'Articles of Impeachment Against Donald John Trump'.
URL: *https://www.congress.gov/116/bills/hres755/BILLS-116hres755enr.pdf*

Costa, P. T., McCrae, R. & Dye, D. A. (1991), 'Domains and facets scales for agreeableness and cconscientiousness: A revision of the NEO personality inventory', *Journal of Personality Assessment* **12**(9), 887–898.

Costa, P. T. & McCrae, R. R. (1992), 'Four ways five factors are basic', *Personality and Individual Differences* **13**(6), 667–673.

D' Adda, G. D., Drouvelis, M. & Nosenzo, D. (2016), 'Norm elicitation in within-subject designs : Testing for order effects', *Journal of Behavioral and Experimental Economics* **62**, 1–7.
URL: *http://dx.doi.org/10.1016/j.socec.2016.02.003*

Dai, Z., Galeotti, F. & Villeval, M. C. (2018), 'Cheating in the lab predicts fraud in the field: An experiment in public transportation', *Management Science* **64**(3), 1081–1100.

Dalton, J. (2019), 'Brexit: Two in give remainers would be upset if their child married leave voter, poll claims'.
URL: *https://www.independent.co.uk/news/uk/politics/brexit-polls-remain-leave-eu-children-marry-a9079756.html*

Danigelis, N. L., Hardy, M. & Cutler, S. J. (2007), 'Population Aging, Intracohort Aging, and Sociopolitical Attitudes', *American Sociological Review* **72**(5), 812–830.

De Geest, L. R. & Stranlund, J. K. (2019), 'Defending public goods and common-pool resources', *Journal of Behavioral and Experimental Economics* **79**(February), 143–154.
URL: *https://doi.org/10.1016/j.socec.2019.02.006*

De Luca, G., Hodler, R., Raschky, P. A. & Valsecchi, M. (2018), 'Ethnic favoritism: An axiom of politics?', *Journal of Development Economics* **132**, 115–129.

DeAndrea, D. C., Carpenter, C., Shulman, H. & Levine, T. R. (2009), 'The relationship between cheating behavior and sensation-seeking', *Personality and Individual Differences* **47**(8), 944–947.
URL: *http://dx.doi.org/10.1016/j.paid.2009.07.021*

Dellavigna, S., List, J. A., Malmendier, U. & Rao, G. (2017), 'Voting to tell others', *Review of Economic Studies* **84**(1), 143–181.

Der Merwe, W. G. V. & Burns, J. (2008), 'What'S in a Name? Racial Identity and Altruism in Post-Apartheid South Africa', *South African Journal of Economics* **76**(2), 266–275.

DeYoung, C. G., Quilty, L. C. & Peterson, J. B. (2007), 'Between Facets and Domains: 10 Aspects of the Big Five', *Journal of Personality and Social Psychology* **93**(5), 880–896.

Deyoung, C. G., Quilty, L. C., Peterson, J. B. & Gray, J. R. (2014), 'Openness to experience, intellect, and cognitive ability', *Journal of Personality Assessment* **96**(1), 46–52.

Di Zheng, J., Schram, A. & Doğan, G. (2020), 'Friend or foe? Social ties in bribery and corruption', *Experimental Economics* (0123456789).

Dimant, E. (2021), 'Hate Trumps Love: The Impact of Political Polarization on Social Preferences', *SSRN Electronic Journal* (May).

Dimant, E., van Kleef, G. A. & Shalvi, S. (2020), 'Requiem for a Nudge: Framing effects in nudging honesty', *Journal of Economic Behavior and Organization* **172**, 247–266.

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

Dreber, A. & Johannesson, M. (2008), 'Gender differences in deception', *Economics Letters* **99**(1), 197–199.

Druckman, J. N., Klar, S., Krupnikov, Y., Levendusky, M. & Ryan, J. B. (2021), 'Affective polarization, local contexts and public opinion in America', *Nature Human Behaviour* **5**(1), 28–38.
URL: *http://dx.doi.org/10.1038/s41562-020-01012-5*

Duclos, J.-y., Esteban, J. & Ray, D. (2003), 'Polarization : Concepts , Measurement , Estimation', *CIRPEE* pp. 1–44.

Dufwenberg, M., Gächter, S. & Hennig-Schmidt, H. (2011), 'The framing of games and the psychology of play', *Games and Economic Behavior* **73**(2), 459–478.
URL: *http://dx.doi.org/10.1016/j.geb.2011.02.003*

Easterly, W. & Levine, R. (1997), 'Africa' s Growth Tragedy : Policies and Ethnic Divisions', *The Quarterly Journal of Economics* **112**(4), 1203–1250.

Eckel, C. C. & Grossman, P. J. (2001), 'Chivalry and solidarity in ultimatum games', *Economic Inquiry* **39**(2), 171–188.

Eckel, C. C. & Grossman, P. J. (2004), 'Giving to secular causes by the religious and nonreligious: An experimental test of the responsiveness of giving to subsidies', *Nonprofit and Voluntary Sector Quarterly* **33**(2), 271–289.

Eckel, C. C. & Grossman, P. J. (2005), 'Managing diversity by creating team identity', *Journal of Economic Behavior and Organization* **58**(3), 371–392.

Eckel, C. C. & Grossman, P. J. (2008), Differences in the Economic Decisions of Men and Women: Experimental Evidence, *in* 'Handbook of Experimental Economics Results', chapter 57, pp. 509–519.

Ellingsen, T. & Östling, R. (2010), 'When Does Communication Improve Coordination?', *The American Economic Review* **100**(4), 1695–1724.

Elster, J. (1989), 'Social norms and economic welfare', *Journal of Economic Perspectives* **3**(4), 99–117.
URL: *http://linkinghub.elsevier.com/retrieve/pii/S0014292197001384*

Erat, S. & Gneezy, U. (2012), 'White Lies', *Management Science* **58**(4), 723–733.

Espinosa, R. & Stoop, J. (2021), 'Do people really want to be informed? Ex-ante evaluations of information-campaign effectiveness', *Experimental Economics* (0123456789).

Esteban, J.-M. & Ray, D. (1994), 'On the Measurement of Poalrization', *Econometrica* **62**(4), 819–851.

Etang, A., Fielding, D. & Knowles, S. (2011), 'Does trust extend beyond the village? Experimental trust and social distance in Cameroon', *Experimental Economics* **14**(1), 15–35.

Evans, A. M. & Revelle, W. (2008), 'Survey and behavioral measurements of interpersonal trust', *Journal of Research in Personality* **42**(6), 1585–1593.
URL: *http://dx.doi.org/10.1016/j.jrp.2008.07.011*

Ewert, B. & Loer, K. (2021), 'Advancing behavioural public policies: in pursuit of a more comprehensive concept', *Policy & Politics* **49**(1), 25–47.

Falk, A. & Zehnder, C. (2013), 'A city-wide experiment on trust discrimination', *Journal of Public Economics* **100**, 15–27.
URL: *http://dx.doi.org/10.1016/j.jpubeco.2013.01.005*

Farrow, K., Grolleau, G. & Ibanez, L. (2017), 'Social Norms and Pro-environmental Behavior: A Review of the Evidence', *Ecological Economics* **140**, 1–13.
URL: *http://dx.doi.org/10.1016/j.ecolecon.2017.04.017*

Fearon, J. D. (2003), 'Ethnic and cultural diversity by country', *Journal of Economic Growth* **8**(2), 195–222.

Fehr, B. E. & Gachter, S. (2000), 'Cooperation and Punishment in Public Goods Experiments', *American Economic Association* **90**(4), 980–994.

Fehr, E. & Fischbacher, U. (2002), 'Why Social Preferences Matter - The Impact of Non-selfish Motives on Competition, Cooperation and Incentives', *The Economic Journal* **112**(478), 1–33.

Fehr, E. & Schurtenberger, I. (2018), 'Normative foundations of human cooperation review-article', *Nature Human Behaviour* **2**(7), 458–468.

Feldhaus, C. & Mans, J. (2014), 'Who do you lie to? Social Identity and the costs of lying', *University of Cologne: Working paper series in Economics (No.76)* pp. 1–28.

Ferraro, P. J. & Cummings, R. G. (2007), 'Cultural diversity, discrimination, and economic outcomes: An experimental analysis', *Economic Inquiry* **45**(2), 217–232.

Fershtman, C., Gneezy, U. & Verboven, F. (2005), 'Discrimination and Nepotism: The Efficiency of the Anonymity Rule', *The Journal of Legal Studies* **34**(2), 371–396.

Fetzer, T. (2019), 'Did austerity cause brexit?', *American Economic Review* **109**(11), 3849–3886.

Fischbacher, U. & Föllmi-Heusi, F. (2013), 'Lies in disguise-an experimental study on cheating', *Journal of the European Economic Association* **11**(3), 525–547.

Fischbacher, U. & Gächter, S. (2010), 'Social preferences, beliefs, and the dynamics of free riding in public goods experiments', *American Economic Review* **100**(1), 541–556.

Fong, C. M. & Luttmer, E. F. (2011), 'Do fairness and race matter in generosity? Evidence from a nationally representative charity experiment', *Journal of Public Economics* **95**(5-6), 372–394.
URL: *http://dx.doi.org/10.1016/j.jpubeco.2010.07.010*

Forsythe, R., Horowitz, J. L., Savin, N. E. & Sefton, M. (1994), 'Fairness in simple bargaining experiments', *Games and Economic Behavior* **6**(3), 347–369.

Fosgaard, T. R., Hansen, L. G. & Piovesan, M. (2013), 'Separating Will from Grace: An experiment on conformity and awareness in cheating', *Journal of Economic Behavior and Organization* **93**, 279–284.
URL: *http://dx.doi.org/10.1016/j.jebo.2013.03.027*

Foster, J. E. & Wolfson, M. C. (2010), 'Polarization and the decline of the middle class: Canada and the U.S.', *Journal of Economic Inequality* **8**(2), 247–273.

Gächter, S. & Fehr, E. (1999), 'Collective action as a social exchange', *Journal of Economic Behavior & Organization* **39**(4), 341–369.

Gächter, S. & Renner, E. (2010), 'The effects of (incentivized) belief elicitation in public goods experiments', *Experimental Economics* **13**(3), 364–377.

Gächter, S. & Schulz, J. F. (2016), 'Intrinsic honesty and the prevalence of rule violations across societies', *Nature* **531**(7595), 496–499.

Gerber, A. S., Huber, G. A., Doherty, D., Dowling, C. M. & Panagopoulos, C. (2013), 'Big Five Personality Traits and Responses to Persuasive Appeals: Results from Voter Turnout Experiments', *Political Behavior* **35**(4), 687–728.

Gerlach, P., Teodorescu, K. & Hertwig, R. (2019), 'The truth about lies: A meta-analysis on dishonest behavior', *Psychological Bulletin* **145**(1), 1–44.

Gerlitz, J.-Y. & Schupp, J. (2005), 'Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale im SOEP', *DIW Research Notes* **4**, 1–44.

Giluk, T. L. & Postlethwaite, B. E. (2015), 'Big Five personality and academic dishonesty: A meta-analytic review', *Personality and Individual Differences* **72**, 59–67.
URL: *http://dx.doi.org/10.1016/j.paid.2014.08.027*

Gino, F. & Ariely, D. (2012), 'The dark side of creativity: Original thinkers can be more dishonest', *Journal of Personality and Social Psychology* **102**(3), 445–459.

Gino, F., Ayal, S. & Ariely, D. (2013), 'Self-Serving Altruism? The Lure of Unethical Actions that Benefit Others', *Journal of Economic Behavior & Organization* **93**(1), 285–292.
URL: *file:///H:/Mis documentos/documentos/separatas/*

Gintis, H., Bowles, S., Boyd, R. & Fehr, E. (2003), 'Explaining altruistic behavior in humans', *Evolution and Human Behavior* **24**(3), 153–172.

Gneezy, U. (2005), 'Deception: The role of consequences', *American Economic Review* **95**(1), 384–394.

Gneezy, U., Kajackaite, A. & Sobel, J. (2018), 'Lying Aversion and the Size of the Lie', *American Economic Review* **108**(2), 419–453.

Gofen, A., Moseley, A., Thomann, E., Weaver, R. K., Gofen, A., Moseley, A., Thomann, E., Weaver, R. K. & Thomann, E. (2021), 'Behavioural governance in the policy process : introduction to the special issue Behavioural governance in the policy process :', *Journal of European Public Policy* **0**(0), 1–25.
URL: *https://doi.org/10.1080/13501763.2021.1912153*

Goldberg, L. R. (1992), 'The Development of Markers for the Big-Five Factor Structure', *Psychological Assessment* **4**(1), 26–42.

Goldberg, L. R. (1993), 'The structure of phenotypic personality traits', *The American Psychologist* **48**(1), 26–34.

Gosling, S. D., Rentfrow, P. J. & Swann, W. B. (2003), 'A very brief measure of the Big-Five personality domains', *Journal of Research in Personality* **37**(6), 504–528.

Graf, R. (2019), Nudging before the nudge? Behavioural traffic safety regulation and the rise of behavioural economics, *in* H. StraSSheim & S. Beck, eds, 'Handbook of Behavioural Change and Public Policy', Cheltenham, UK: Massachusetts, USA, pp. 23–37.

Graham, J., Haidt, J. & Nosek, B. A. (2009), 'Liberals and Conservatives Rely on Different Sets of Moral Foundations', *Journal of Personality and Social Psychology* **96**(5), 1029–1046.

Graziano, W. G., Hair, E. C. & Finch, J. F. (1997), 'Competitiveness mediates the link between personality and group performance', *Journal of Personality and Social Psychology* **73**(6), 1394–1408.

Green, D. (2012), Shifts in Social Norms Often Underpin Change, *in* 'How Change Happens', number 2016, Oxford University Press, Oxford, chapter 3, pp. 47–68.

Greiner, B. (2015), 'Subject pool recruitment procedures: organizing experiments with ORSEE', *Journal of the Economic Science Association* **1**(1), 114–125.

Grolleau, G., Kocher, M. G. & Sutan, A. (2016), 'Cheating and loss aversion : do people lie more to avoid a loss ?', *Management Science* **62**(12), 3428–3438.

Grosskopf, B. & Pearce, G. (2017), 'Discrimination in a deprived neighbourhood: An artefactual field experiment', *Journal of Economic Behavior and Organization* **141**, 29–42.
URL: *http://dx.doi.org/10.1016/j.jebo.2017.05.016*

Guala, F., Mittone, L. & Ploner, M. (2013), 'Group membership, team preferences, and expectations', *Journal of Economic Behavior and Organization* **86**, 183–190.
URL: *http://dx.doi.org/10.1016/j.jebo.2012.12.003*

Guryan, J. & Charles, K. K. (2013), 'Taste-based or statistical discrimination: The economics of discrimination returns to its roots', *Economic Journal* **123**(572), 417–432.

Güth, W., Ockenfels, P. & Wendel, M. (1997), 'Cooperation based on trust: An experimental investigation', *Journal of Economic Psychology* **18**(1), 15–43.

Haagsma, R. & Mouche, P. V. (2013), 'Egalitarian norms, economic development, and ethnic polarization', *Journal of Comparative Economics* **41**(3), 719–744.
URL: *http://dx.doi.org/10.1016/j.jce.2012.10.005*

Hahn, E., Gottschling, J. & Spinath, F. M. (2012), 'Short measurements of personality - Validity and reliability of the GSOEP Big Five Inventory (BFI-S)', *Journal of Research in Personality* **46**(3), 355–359.
URL: *http://dx.doi.org/10.1016/j.jrp.2012.03.008*

Hanel, P. H. & Wolf, L. J. (2020), 'Leavers and Remainers after the Brexit referendum: More united than divided after all?', *British Journal of Social Psychology* **59**(2), 470–493.

Hansen, P. G. & Jespersen, A. M. (2013), 'Nudge and the Manipulation of Choice: A Framework for the Responsible Use of the Nudge Approach to Behaviour Change in Public Policy', *European Journal of Risk Regulation* **4**(1), 3–28.
URL: *http://www.jstor.org.uplib.idm.oclc.org/stable/24323381*

Harris, D., Herrmann, B., Kontoleon, A. & Newton, J. (2015), 'Is it a norm to favour your own group?', *Experimental Economics* **18**(3).

Hart, C. L., Lemon, R., Curtis, D. A. & Griffith, J. D. (2020), 'Personality Traits Associated with Various Forms of Lying', *Psychological Studies* **65**(3), 239–246.
URL: *https://doi.org/10.1007/s12646-020-00563-x*

Hattenstone, S. (2020), 'In Boris Johnson's long history of lies, the Marcus Rashford one is the strangest', *Guardian* .
URL: *https://www.theguardian.com/commentisfree/2020/jun/21/boris-johnson-lies-marcus-rashford-prime-minister*

Hayes, N. & Joseph, S. (2003), 'Big 5 correlates of three measures of subjective well-being', *Personality and Individual Differences* **34**(4), 723–727.

Heinicke, F., König-Kersting, C. & Schmidt, R. (2022), 'Injunctive vs. descriptive social norms and reference group dependence', *Journal of Economic Behavior and Organization* **195**, 199–218.
URL: *https://doi.org/10.1016/j.jebo.2022.01.008*

Hendriks, C. M., Ercan, S. A. & Duus, S. (2019), 'Listening in polarised controversies: a study of listening practices in the public sphere', *Policy Sciences* **52**(1), 137–151.
URL: *https://doi.org/10.1007/s11077-018-9343-3*

Hendy, N. T. & Montargot, N. (2019), 'Understanding academic dishonesty among business school students in france using the theory of planned behavior', *The International Journal of Management Education* **17**(1), 85–93.
URL: *https://www.sciencedirect.com/science/article/pii/S1472811718302258*

Henrich, J., Ensminger, J., Mcelreath, R., Barr, A., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., Ziker, J., Henrich, J., Ensminger, J., Mcelreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D. & Ziker, J. (2016), 'Markets , Religion , Community Size , and the Evolution of Fairness and Punishment', *Science* **327**(5972), 1480–1484.

Henrich, J., Heine, S. J. & Norenzayan, A. (2010), 'The weirdest people in the world?', *Behavioral and Brain Sciences* **33**(2-3), 61–83.

Hertwig, R. & Ortmann, A. (2008b), 'Deception in social psychological experiments: Two misconceptions and a research agenda', *Social Psychology Quarterly* **71**(3), 222–227.

Hilbig, B. E., Zettler, I. & Heydasch, T. (2012), 'Personality, Punishment and Public Goods: Strategic Shifts Towards Cooperation as a Matter of Dispositional Honesty-Humility', *European Journal of Personality* **26**(3), 245–254.

Hirsh, J. B., DeYoung, C. G., Xu, X. & Peterson, J. B. (2010), 'Compassionate liberals and polite conservatives: Associations of agreeableness with political ideology and moral values', *Personality and Social Psychology Bulletin* **36**(5), 655–664.

Hirsh, J. B. & Peterson, J. B. (2009), 'Extraversion, neuroticism, and the prisoner's dilemma', *Personality and Individual Differences* **46**(2), 254–256.
URL: *http://dx.doi.org/10.1016/j.paid.2008.10.006*

Hobolt, S. B. (2016), 'The Brexit vote: a divided nation, a divided continent', *Journal of European Public Policy* **23**(9), 1259–1277.
URL: *http://dx.doi.org/10.1080/13501763.2016.1225785*

Hobolt, S. B., Leeper, T. J. & Tilley, J. (2021), 'Divided by the Vote: Affective Polarization in the Wake of the Brexit Referendum', *British Journal of Political Science* **51**(4), 1476–1493.

Holzmeister, F., Huber, J., Kirchler, M. & Schwaiger, R. (2022), 'Nudging Debtors to Pay Their Debt: Two Randomized Controlled Trials', *Journal of Economic Behavior and Organization* **198**, 535–551.
URL: *https://doi.org/10.1016/j.jebo.2022.04.006*

Hong, F. Y., Chiu, S. I. & Huang, D. H. (2012), 'A model of the relationship between psychological characteristics, mobile phone addiction and use of mobile phones by Taiwanese university female students', *Computers in Human Behavior* **28**(6), 2152–2159.
URL: *http://dx.doi.org/10.1016/j.chb.2012.06.020*

Hornung, J., Bandelow, N. C. & Vogeler, C. S. (2019), 'Social identities in the policy process', *Policy Sciences* **52**(2), 211–231.
URL: *https://doi.org/10.1007/s11077-018-9340-6*

Houser, D., List, J. A., Piovesan, M., Samek, A. & Winter, J. (2016), 'Dishonesty: From parents to children', *European Economic Review* **82**, 242–254.

Howlett, M. & Leong, C. (2022), 'What is Behavioral in Policy Studies?', *Journal of Behavioral Public Administration* **5**(1).

Howlett, M., Ramesh, M. & Capano, G. (2020), 'Policy-Makers, Policy-Takers and Policy Tools: Dealing with Behaviourial Issues in Policy Design', *Journal of Comparative Policy Analysis: Research and Practice* **22**(6), 487–497.
URL: *https://doi.org/10.1080/13876988.2020.1774367*

Hugh-Jones, D. (2016), 'Honesty, beliefs about honesty, and economic growth in 15 countries', *Journal of Economic Behavior and Organization* **127**, 99–114.
URL: *http://dx.doi.org/10.1016/j.jebo.2016.04.012*

Huitema, D., Jordan, A., Munaretto, S. & Hildén, M. (2018), 'Policy experimentation: core concepts, political dynamics, governance and impacts', *Policy Sciences* **51**(2), 143–159.
URL: *https://doi.org/10.1007/s11077-018-9321-9*

Hummel, D. & Maedche, A. (2019), 'How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies', *Journal of Behavioral and Experimental Economics* **80**(February), 47–58.

Hydock, C. (2018), 'Assessing and overcoming participant dishonesty in online data collection', *Behavior Research Methods* **50**(4), 1563–1567.

Iriberri, N. & Rey-Biel, P. (2021), 'Brave Boys and Play-it-Safe Girls: Gender Differences in Willingness to Guess in a Large Scale Natural Field Experiment', *European Economic Review* **131**, 103603.
URL: *https://doi.org/10.1016/j.euroecorev.2020.103603*

Isler, O. & Gächter, S. (2022), 'Conforming with peers in honesty and cooperation', *Journal of Economic Behavior & Organization* **195**, 75–86.
URL: *https://doi.org/10.1016/j.jebo.2021.12.026*

Isler, O., Yilmaz, O. & John Maule, A. (2021), 'Religion, parochialism and intuitive cooperation', *Nature Human Behaviour* **5**(4), 512–521.
URL: *http://dx.doi.org/10.1038/s41562-020-01014-3*

Iyengar, S. & Westwood, S. J. (2015), 'Fear and Loathing across Party Lines: New Evidence on Group Polarization', *American Journal of Political Science* **59**(3), 690–707.

Jacobsen, C., Fosgaard, T. R. & Pascual-Ezama, D. (2018), 'Why Do We Lie? a Practical Guide To the Dishonesty Literature', *Journal of Economic Surveys* **32**(2), 357–387.

Jahedi, S. & Méndez, F. (2014), 'On the advantages and disadvantages of subjective measures', *Journal of Economic Behavior and Organization* **98**, 97–114.
URL: *http://dx.doi.org/10.1016/j.jebo.2013.12.016*

Jiang, T. (2013), 'Cheating in mind games: The subtlety of rules matters', *Journal of Economic Behavior and Organization* **93**, 328–336.
URL: *http://dx.doi.org/10.1016/j.jebo.2013.04.003*

John, O. P., Donahue, E. M. & Kentle, R. L. (1991), 'Big five inventory', *Journal of Personality and Social Psychology* .

John, O. P., Naumann, L. P. & Soto, C. J. (2008), Paradigm shift to the integrative Big Five Trait taxonomy: History, measurement, and conceptual issues, *in* 'Handbook of personality: Theory and research', pp. 114–158.

John, O. P. & Srivastava, S. (1999), The Big Five trait taxonomy: History, measurement, and theoretical perspectives, *in* L. Pervin & O. P. John, eds, 'Handbook of personality: Theory and research', 2nd edn, number 510, pp. 1–71.

John, P., Smith, G. & Stoker, G. (2009), 'Nudge Nudge, Think Think: Two Strategies for Changing Civic Behaviour', *The Political Quartely* **80**(3), 361–370.

Jost, J. T., Glaser, J., Kruglanski, A. W. & Sulloway, F. J. (2003), 'Political Conservatism as Motivated Social Cognition', *Psychological Bulletin* **129**(3), 339–375.

Kagel, J. & McGee, P. (2014), 'Personality and cooperation in finitely repeated prisoner's dilemma games', *Economics Letters* **124**(2), 274–277.
URL: *http://dx.doi.org/10.1016/j.econlet.2014.05.034*

Kajackaite, A. (2018), 'Lying about luck versus lying about performance', *Journal of Economic Behavior and Organization* **153**, 194–199.
URL: *https://doi.org/10.1016/j.jebo.2018.07.010*

Kajackaite, A. & Gneezy, U. (2017), 'Incentives and cheating', *Games and Economic Behavior* **102**, 433–444.
URL: *http://dx.doi.org/10.1016/j.geb.2017.01.015*

Kallgren, C. A., Reno, R. R. & Cialdini, R. B. (2000), 'A focus theory of normative conduct: When norms do and do not affect behavior', *Personality and Social Psychology Bulletin* **26**(8), 1002–1012.

Kant, I. (1949), On a Supported Right to Lie from Altruistic Motives, *in* L. Beck White, ed., 'Critique of Practical Reason and Other Writings in Moral Philosophy', pp. 346–350.

Karlan, D. S. (2005), 'Using Experimental Economics to Measure Social Capital and Predict Financial Decisions', *American Economic Review* **95**(5), 1688–1699.

Kartik, N., Tercieux, O. & Holden, R. (2014), 'Simple mechanisms and preferences for honesty', *Games and Economic Behavior* **83**, 284–290.
URL: *http://dx.doi.org/10.1016/j.geb.2013.11.011*

Khalmetski, K. & Sliwka, D. (2019), 'Disguising lies - Image Concerns and Partial Lying in Cheating Games', *American Economic Journal: Microeconomics* **11**(4), 79–110.

Kim, B. Y., Choi, S., Lee, J., Lee, S. & Choi, K. (2017), 'Do Institutions Affect Social Preferences? Evidence from Divided Korea', *Journal of Comparative Economics* **45**(4), 865–888.

Kimbrough, E. O. & Vostroknutov, A. (2016), 'Norms Make Preferences Social', *Journal of the European Economic Association* **14**(3), 608–638.

Klein, S. A., Heck, D. W., Reese, G. & Hilbig, B. E. (2019), 'On the relationship between Openness to Experience, political orientation, and pro-environmental behavior', *Personality and Individual Differences* **138**(July 2018), 344–348.
URL: *https://doi.org/10.1016/j.paid.2018.10.017*

Kline, R., Bankert, A., Levitan, L. & Kraft, P. (2019), 'Personality and prosocial behavior: A multilevel meta-analysis', *Political Science Research and Methods* **7**(1), 125–142.

Kocher, M. G., Schudy, S. & Spantig, L. (2018), 'I lie? We lie! Why? Experimental evidence on a dishonesty shift in groups', *Management Science* **64**(9), 3995–4008.

Kohler, S. (2013), 'More Fair Play in an Ultimatum Game after Resettlement in Zimbabwe: A Field Experiment and a Structural Model', *PLoS ONE* **8**(5).

Komaromi, P. (2016), Post-referendum Racism and Xenophobia: The Role of Social Media Activism in Challenging the Normalisation of Xeno-racist Narratives, Technical report, Institute of race relations.
URL: *https://irr.org.uk/article/post-referendum-racism-and-the-importance-of-social-activism/*

Krupka, E. L. & Weber, R. A. (2013), 'Identifying social norms using coordination games: Why does dictator game sharing vary?', *Journal of the European Economic Association* **11**(3), 495–524.

Kuper, S. (2022), 'US polarisation is a unique phenomenon', *Financial Times* .

Laakso, L. (2002), 'The politics of international election observation: The case of Zimbabwe in 2000', *Journal of Modern African Studies* **40**(3), 437–464.

Lamont, M. & Molnár, V. (2002), 'The Study of Boundaries in the Social Sciences', *Annual Review of Sociology* **28**(1), 167–195.

Lane, T. (2016), 'Discrimination in the laboratory: A meta-analysis of economics experiments', *European Economic Review* **90**, 375–402.
URL: *http://dx.doi.org/10.1016/j.euroecorev.2015.11.011*

Lang, F. R., John, D., Lüdtke, O., Schupp, J. & Wagner, G. G. (2011), 'Short assessment of the Big Five: Robust across survey methods except telephone interviewing', *Behavior Research Methods* **43**(2), 548–567.

Lang, F. R., Lüdtke, O. & Asendorpf, J. B. (2001), 'Testgüte und psychometrische Äquivalenz der deutschen Version des Big Five Inventory (BFI) bei jungen, mittelalten und alten Erwachsenen', *Diagnostica* **47**(3), 111–121.

Lawn, E. C. R., Zhao, K., Laham, S. M. & Smillie, L. D. (2021), 'Prosociality Beyond Big Five Agreeableness and HEXACO Honesty-Humility: Is Openness/Intellect Associated With Cooperativeness in the Public Goods Game?', *European Journal of Personality* p. 089020702110281.

Lawton, C. & Ackrill, R. (2016), 'Hard Evidence: how ares with low immigration voted mainly Brexit'.
URL: *https://theconversation.com/hard-evidence-how-areas-with-low-immigration-voted-mainly-for-brexit-62138*

LeBas, A. (2018), 'Can Polarization Be Positive? Conflict and Institutional Development in Africa', *American Behavioral Scientist* **62**(1), 59–74.

LeBas, A. & Munemo, N. (2019), 'Elite Conflict, Compromise, and Enduring Authoritarianism: Polarization in Zimbabwe, 1980–2008', *Annals of the American Academy of Political and Social Science* **681**(1), 209–226.

Levy, D. A., Aslan, B. & Bironzo, D. (2016), UK PRESS COVERAGE OF THE EU REFERENDUM, Technical report.

Liberman, V., Samuels, S. M. & Ross, L. (2004), 'The name of the game: Predictive power of reputations versus situational labels in determining Prisoner's Dilemma game moves', *Personality and Social Psychology Bulletin* **30**(9), 1175–1185.

Lodge, M. & Wegrich, K. (2016), 'The Rationality Paradox of Nudge: Rational Tools of Government in a World of Bounded Rationality', *Law and Policy* **38**(3), 250–267.

Lorenz-spreen, P., Lewandowsky, S., Sunstein, C. R. & Hertwig, R. (2020), 'How behavioural sciences can promote truth, autonomy and democratic discourse online', *Nature Human Behaviour* **4**(11), 1102–1109.
URL: *http://dx.doi.org/10.1038/s41562-020-0889-7*

Lu, L. & Argyle, M. (1991), 'Happiness and cooperation', *Personality and Individual Differences* **12**(10), 1019–1030.

Luhan, W. J., Kocher, M. G. & Sutter, M. (2009), 'Group polarization in the team dictator game reconsidered', *Experimental Economics* **12**(1), 26–41.

MacInnis, C. C., Boss, H. C. & Bourdage, J. S. (2020), 'More evidence of participant misrepresentation on Mturk and investigating who misrepresents', *Personality and Individual Differences* **152**(August 2019), 109603.
URL: *https://doi.org/10.1016/j.paid.2019.109603*

Mackie, D. M. (1986), 'Social Identification Effects in Group Polarization', *Journal of Personality and Social Psychology* **50**(4), 720–728.

Macy, M., Deri, S., Ruch, A. & Tong, N. (2019), 'Opinion cascades and the unpredictability of partisan polarization', *Science Advances* **5**(8), 1–8.

Mark, G. & Ganzach, Y. (2014), 'Personality and Internet usage: A large-scale representative study of young adults', *Computers in Human Behavior* **36**, 274–281.
URL: *http://dx.doi.org/10.1016/j.chb.2014.03.060*

Matsushima, H. (2008), 'Role of honesty in full implementation', *Journal of Economic Theory* **139**(1), 353–359.

Mazar, N., Amir, O. & Ariely, D. (2008), 'The dishonesty of honest people: A theory of self-concept maintenance', *Journal of Marketing Research* **45**(6), 633–644.

McCoy, J., Rahman, T. & Somer, M. (2018), 'Polarization and the Global Crisis of Democracy: Common Patterns, Dynamics, and Pernicious Consequences for Democratic Polities', *American Behavioral Scientist* **62**(1), 16–42.

McCrae, R. R. & Costa, P. T. (1997), 'Personality Trait Structures as a Human Universal', *American Psychologist* **52**(5), 509–516.

McTernan, M., Love, P. & Rettinger, D. (2014), 'The Influence of Personality on the Decision to Cheat', *Ethics and Behavior* **24**(1), 53–72.

Mill, W. & Morgan, J. (2021), 'The cost of a divided America: an experimental study into destructive behavior', *Experimental Economics* (0123456789).
URL: *https://doi.org/10.1007/s10683-021-09737-4*

Miller, D. T. & McFarland, C. (1987), 'Pluralistic Ignorance: When Similarity is Interpreted as Dissimilarity', *Journal of Personality and Social Psychology* **53**(2), 298–305.

Miller, D. T. & Prentice, D. A. (2016), 'Changing norms to change behavior', *Annual Review of Psychology* **67**(July 2015), 339–361.

Miller, J. D. & Lynam, D. (2001), 'Structural models of personality and their relation to antisocial behavior: A meta-analytic review', *Criminology* **39**(4), 765–798.

Mlambo, A. S. (2013), 'Becoming Zimbabwe or Becoming Zimbabwean: Identity, Nationalism and State-building', *Africa Spectrum* **48**(1), 49–70.

Mol, J., Botzen, W., Blasch, J., Kranzler, E. & Kunreuther, H. C. (2020), 'All by Myself? Testing Descriptive Social Norm-Nudges to Increase Flood Preparedness among Homeowners', *SSRN Electronic Journal* (November).

Monroe, K. R., Hankin, J. & Van Vechten, R. B. (2000), 'The Pyschological Foundations of Identity Politics', *Annual Review of Political Science* **3**(1), 419–447.

Montalvo, J. G. & Reynal-Querol, M. (2005), 'Ethnic Polarization, Potential Conflict, and Civil Wars', *American Economic Review* **95**(3), 796–816.

Mounk, Y. (2018), *The People vs. Democracy: Why our freedom is in danger and how to save it*, Havard University Press, Cambridge, Massachusetts; Lond, England.

Mpofu, S. (2013), 'Social media and the politics of ethnicity in Zimbabwe', *Ecquid Novi: African Journalism Studies* **34**(1), 115–122.

Murphy, P. R. (2012), 'Attitude, Machiavellianism and the rationalization of misreporting', *Accounting, Organizations and Society* **37**(4), 242–259.
URL: *http://dx.doi.org/10.1016/j.aos.2012.04.002*

Muzondidya, J. & Ndlovu-Gatsheni, S. (2007), '' Echoing Silences ': Ethnicity in post-colonial Zimbabwe, 1980-2007', *African and Asian Studies* **7**(2), 275–297.

Nagatsu, M. (2015), 'Social Nudges: Their Mechanisms and Justification', *Review of Philosophy and Psychology* **6**(3), 481–494.

Nannicini, T., Stella, A., Tabellini, G. & Troiano, U. (2013), 'Social capital and political accountability', *American Economic Journal: Economic Policy* **5**(2), 222–250.

Ndhlovu, F. (2007), 'The role of discourse in identity formation and the manufacture of ethnic minorities in Zimbabwe', *Journal of Multicultural Discourses* **2**(2), 131–147.

Ndlovu-Gatsheni, S. (2011), 'Nation Building in Zimbabwe and the Challenges of Ndebele Particularism', *African Journal on Conflict Resolution* **8**(3), 27–55.

Ndlovu-Gatsheni, S. J. (2012), 'The death of the Subject with a capital S and the perils of belonging: A study of the construction of ethnocracy in Zimbabwe', *Critical Arts* **26**(4), 525–546.

Ockenfels, A. & Werner, P. (2014), 'Beliefs and ingroup favoritism', *Journal of Economic Behavior and Organization* **108**, 453–462.
URL: *http://dx.doi.org/10.1016/j.jebo.2013.12.003*

Oh, I. S., Lee, K., Ashton, M. C. & de Vries, R. E. (2011), 'Are dishonest extraverts more harmful than dishonest introverts? The interaction effects of honesty-humility and extraversion in predicting workplace deviance', *Applied Psychology* **60**(3), 496–516.

Ohtomo, S. & Hirose, Y. (2007), 'The dual-process of reactive and intentional decision-making involved in eco-friendly behavior', *Journal of Environmental Psychology* **27**(2), 117–125.

Ölander, F. & Thøgersen, J. (2014), 'Informing Versus Nudging in Environmental Policy', *Journal of Consumer Policy* **37**(3), 341–356.

Onslow, S. (2011), 'Zimbabwe and Political Transition Zimbabwe and', *The London School of Economics and Political Science* pp. 1–16.

Ostrom, E. (1998), 'A Behavioral Approach to the Rational Choice Theory of Collective Action: Presidential Address, American Political Science Association, 1997', *American Political Science Review* **92**(1), 1–22.

Ostrom, E. (2014), 'Collective action and the evolution of social norms', *Journal of Natural Resources Policy Research* **6**(4), 235–252.

Palan, S. & Schitter, C. (2018), 'Prolific.ac—A subject pool for online experiments', *Journal of Behavioral and Experimental Finance* **17**, 22–27.
URL: *https://doi.org/10.1016/j.jbef.2017.12.004*

Parks-Leduc, L., Feldman, G. & Bardi, A. (2015), 'Personality Traits and Personal Values: A Meta-Analysis', *Personality and Social Psychology Review* **19**(1), 3–29.

Paunonen, S. V. & Ashton, M. C. (2001), 'Big five factors and facets and the prediction of behavior', *Journal of Personality and Social Psychology* **81**(3), 524–539.

Pavot, W., Diener, E. & Fujita, F. (1990), 'Extraversion and happiness', *Personality and Individual Differences* **11**(12), 1299–1306.

Phelps, E. S. (1972), 'The Statistical Theory of Racism and Sexism', *The American Economic Review* **62**(4), 659–661.
URL: *https://www-jstor-org.up.idm.oclc.org/stable/pdf/1806107.pdf?ab_segments=0%2FSYC-5878%2Fcontrol&refreqid=fastly-default%3Ada5733f1b0df1dd6216a4afa0534adc3*

Posner, D. N. (2005), *Institutions and Ethnic Politics in Africa*, Political Economy of Institutions and Decisions, Cambridge University Press.

Qualtrics (2020), 'Qualtrics'.
URL: *https://www.qualtrics.com*

Rammstedt, B. & John, O. P. (2007), 'Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German', *Journal of Research in Personality* **41**(1), 203–212.

Rangel, A., Camerer, C. & Montague, P. R. (2008), 'A framework for studying the neurobiology of value-based decision making', *Nature Reviews Neuroscience* **9**(7), 545–556.

Rege, Mari; Telle, K. (2004), 'The impact of social approval and framing on cooperation in public good situations', *Journal of Public Economics* **88**(7), 1625–1644.

Rentfrow, P. J., Jost, J. T., Gosling, S. D. & Potter, J. (2009), Statewide Differences in Personality Predict Voting Patterns in 1996-2004 U.S. Presidential Elections, *in* J. T. Jost, A. C. Kay & H. Thorisdottir, eds, 'Social and Psychological Bases of Ideology and System Justification', Oxford University Press, New York, pp. 314–348.

Reuben, E. & Suetens, S. (2012), 'Revisiting strategic versus non-strategic cooperation', *Experimental Economics* **15**(1), 24–43.

Robalo, P., Schram, A. & Sonnemans, J. (2017), 'Other-regarding preferences, in-group bias and political participation: An experiment', *Journal of Economic Psychology* **62**, 130–154.
URL: *http://dx.doi.org/10.1016/j.joep.2017.04.009*

Roberts, B. W., Lejuez, C., Krueger, R. F., Richards, J. M. & Hill, P. L. (2014), 'What is conscientiousness and how can it be assessed?', *Developmental Psychology* **50**(5), 1315–1330.

Robinson, A. L. (2016), 'Nationalism and Ethnic-Based Trust: Evidence From an African Border Region', *Comparative Political Studies* **49**(14), 1819–1854.
URL: *https://doi.org/10.1177/0010414016628269*

Roth, A. E. & Erev, I. (1995), 'Learning in Extensive-Form Games : Experimental Data and Simple Dynamic Models in the Intermediate Term', *Games and Economic Behavior* **8**(1), 164–212.

Rustichini, A., DeYoung, C. G., Anderson, J. E. & Burks, S. V. (2016), 'Toward the integration of personality theory and decision theory in explaining economic behavior: An experimental investigation', *Journal of Behavioral and Experimental Economics* **64**, 122–137.
URL: *http://dx.doi.org/10.1016/j.socec.2016.04.019*

Ryvkin, D., Serra, D. & Tremewan, J. (2017), 'I paid a bribe: An experiment on information sharing and extortionary corruption', *European Economic Review* **94**, 1–22.

Sabater-Grande, G., García-Gallego, A., Georgantzís, N. & Herranz-Zarzoso, N. (2022), 'The effects of personality, risk and other-regarding attitudes on trust and reciprocity', *Journal of Behavioral and Experimental Economics* **96**, 101797.
URL: *https://doi.org/10.1016/j.socec.2021.101797*

Schaub, M. (2017), 'Second-order ethnic diversity: The spatial pattern of diversity, competition and cooperation in Africa', *Political Geography* **59**, 103–116.
URL: *http://dx.doi.org/10.1016/j.polgeo.2017.01.006*

Schotter, A. & Trevino, I. (2014), 'Belief elicitation in the laboratory', *Annual Review of Economics* **6**, 103–128.

Shaffer, J. A. & Postlethwaite, B. E. (2012), 'A matter of context: A meta-analytic investigation of the relative validity of contextualized and noncontextualized personality measures', *Personnel Psychology* **65**(3), 445–494.

Shalvi, S., Handgraaf, M. J. & De Dreu, C. K. (2011), 'People avoid situations that enable them to deceive others', *Journal of Experimental Social Psychology* **47**(6), 1096–1106.
URL: *http://dx.doi.org/10.1016/j.jesp.2011.04.015*

Shalvi, S. & Leiser, D. (2013), 'Moral firmness', *Journal of Economic Behavior and Organization* **93**, 400–407.
URL: *http://dx.doi.org/10.1016/j.jebo.2013.03.014*

Shu, L. L., Gino, F. & Bazerman, M. H. (2011), 'Dishonest deed, clear conscience: When cheating leads to moral disengagement and motivated forgetting', *Personality and Social Psychology Bulletin* **37**(3), 330–349.

Sibley, C. G., Osborne, D. & Duckitt, J. (2012), 'Personality and political orientation: Meta-analysis and test of a Threat-Constraint Model', *Journal of Research in Personality* **46**(6), 664–677.
URL: *http://dx.doi.org/10.1016/j.jrp.2012.08.002*

Slonim, R. & Guillen, P. (2010), 'Gender selection discrimination: Evidence from a Trust game', *Journal of Economic Behavior and Organization* **76**(2), 385–405.
URL: *http://dx.doi.org/10.1016/j.jebo.2010.06.016*

Smerdon, D., Offerman, T. & Gneezy, U. (2020), ''Everybody's doing it': on the persistence of bad social norms', *Experimental Economics* **23**, 392–420.
URL: *https://doi.org/10.1007/s10683-019-09616-z*

Snowberg, E. & Yariv, L. (2021), 'Testing the Waters: Behavior across Participant Pools', *American Economic Review* **111**(2), 687–719.

Solaz, H., De Vries, C. E. & de Geus, R. A. (2019), 'In-Group Loyalty and the Punishment of Corruption', *Comparative Political Studies* **52**(6), 896–926.

Su, S. (2022), 'Updating politicized beliefs: How motivated reasoning contributes to polarization', *Journal of Behavioral and Experimental Economics* **96**(October 2020), 101799.
URL: *https://doi.org/10.1016/j.socec.2021.101799*

Sunstein, C. (1999), 'The Law of Group Polarization', *University of Chicago Law School, John M. Olin Law & Economics Working Paper No.91* pp. 80–101.

Sunstein, C. R. (2017), 'Nudges that fail', *Behavioural Public Policy* **1**(1), 4–25.

Sunstein, C. R. (2019*a*), *Conformity: The Power of Social Influences*, New York University Press.
URL: *https://doi.org/10.18574/9781479896585*

Sunstein, C. R. (2019*b*), Which nudges do people like? A national survey, *in* H. StraSSheim & S. Beck, eds, 'Handbook of Behavioural Change and Public Policy', Edward Elgar, Cheltenham, UK; Massachusetts, USA, pp. 285–303.

Sunstein, C. R. (2021), 'On the Evaluation of Behaviorally Informed Interventions', *SSRN Electronic Journal* .

Sunstein, C. R., Reisch, L. A. & Rauber, J. (2018), 'A worldwide consensus on nudging? Not quite, but almost', *Regulation and Governance* **12**(1), 3–22.

Sutter, M. (2019), 'Deception through Telling the Truth ?! Experimental Evidence from Individuals and Teams', *The Economic Journal* **119**(534), 47–60.

Swope, K. J., Cadigan, J., Schmitt, P. M. & Shupp, R. (2008), 'Personality preferences in laboratory economics experiments', *Journal of Socio-Economics* **37**(3), 998–1009.

Taijfel, H. (1970), 'Experiments in intergroup discrimination.', *Scientific American* **223**(5), 96–102.

Tajfel, H., Billig, M., Bundy, R. & Flament, C. (1971), 'Social categorization and intergroup behaviour', *European Journal of Social Psychology* **1**(2), 149–178.

Talisse, R. (2022), 'Not all polarization is bad, but the US could be in trouble'.
URL: *https://theconversation.com/not-all-polarization-is-bad-but-the-us-could-be-in-trouble-173833*

Tanaka, T. & Camerer, C. F. (2016), 'Trait perceptions influence economic out-group bias: lab and field evidence from Vietnam', *Experimental Economics* **19**(3), 513–534.

te Velde, V. L. & Louis, W. (2022), 'Conformity to descriptive norms', *Journal of Economic Behavior and Organization* **200**, 204–222.

Teney, C., Lacewell, O. P. & De Wilde, P. (2013), 'Winners and losers of globalization in Europe: Attitudes and ideologies', *European Political Science Review* **6**(4), 575–595.

Thaler, R. H. (2018), 'From cashews to nudges: The evolution of behavioral economics', *American Economic Review* **108**(6), 1265–1287.

Thaler, R. H. & Sunstein, C. R. (2008), *Nudge: Improving decisions about health, wealth, and happiness.*, Yale University Press, New Haven, CT, US.

Thaler, R. H. & Sunstein, C. R. (2009), *Nudge : improving decisions about health, wealth and happiness*, Penguin Books.

The Electoral Commission (2017), 'Results and turnout at the EU referendum'.
URL: *https://www.electoralcommission.org.uk/who-we-are-and-what-we-do/elections-and-referendums/past-elections-and-referendums/eu-referendum/results-and-turnout-eu-referendum*

Thielmann, I., Spadaro, G. & Balliet, D. (2020), 'Personality and prosocial behavior: A theoretical framework and meta-analysis', *Psychological Bulletin* **146**(1), 30–90.

Thomsson, K. M. & Vostroknutov, A. (2017), 'Small-world conservatives and rigid liberals: Attitudes towards sharing in self-proclaimed left and right', *Journal of Economic Behavior and Organization* **135**, 181–192.
URL: *http://dx.doi.org/10.1016/j.jebo.2017.01.023*

Tilly, C. (2004), 'Social boundary mechanisms', *Philosophy of the Social Sciences* **34**(2), 211–236.

Toma, C. L., Hancock, J. T. & Ellison, N. B. (2008), 'Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles', *Personality and Social Psychology Bulletin* **34**(8), 1023–1036.

Tracy, J. L. & Robins, R. W. (2007), 'The psychological structure of pride: A tale of two facets', *Journal of Personality and Social Psychology* **92**(3), 506–525.

Turner, J. C., Brown, R. J. & Tajfel, H. (1979), 'Social comparison and group interest in ingroup favouritism', *European Journal of Social Psychology* **9**(2), 187–204.

Tyers, R. (2019), Macro libertarianism and micro paternalism: goverance in an age of nudging, *in* H. StraSSheim & S. Beck, eds, 'Handbook of Behavioural Change and Public Policy', Edward Elgar, Cheltenham, UK; Massachusetts, USA, pp. 332–343.

Virdee, S. & McGeever, B. (2018), 'Racism, Crisis, Brexit', *Ethnic and Racial Studies* **41**(10), 1802–1819.

Wang, X. & Navarro-Martinez, D. (2023), 'Increasing the external validity of social preference games by reducing measurement error', *Games and Economic Behavior* **141**, 261–285.
URL: *https://doi.org/10.1016/j.geb.2023.06.006*

Weisel, O. & Shalvi, S. (2015), 'The collaborative roots of corruption', *Proceedings of the National Academy of Sciences of the United States of America* **112**(34), 10651–10656.

Wessling, K. S., Huber, J. & Netzer, O. (2017), 'MTurk character misrepresentation: Assessment and solutions', *Journal of Consumer Research* **44**(1), 211–230.

Wiltermuth, S. S. (2011), 'Cheating more when the spoils are split', *Organizational Behavior and Human Decision Processes* **115**(2), 157–168.
URL: *https://www.sciencedirect.com/science/article/pii/S0749597810000841*

Yamagishi, T. & Kiyonari, T. (2000), 'The Group as the Container of Generalized Reciprocity', *Social Psychology Quarterly* **63**(2), 116–132.

Zhang, X. & Kanbur, R. (2009), 'What difference do polarization measures make?: An application to China', *Regional Inequality in China: Trends, Explanations and Policy Responses* **0388**, 31–44.

Zimbabwe Human Rights NGO Forum (2001), Enforcing The Rule Of Law in Zimabwe, Technical Report 1.

A

Appendix A for Chapter 2

Appendix A.1 - Extension to Theoretical framework

Let following the norm give no disutility to player i (the sender). Choosing to defect or keep everything leads to disutility. The norm-dependent variable γ measures the magnitude. We posit that the magnitude of the norm will change depending on the treatment. Take, for instance, the Prisoner's Dilemma, the norm prescribes the most equitable outcome (B,B) or (C,C). However, the dominant strategy, which is consistent with the rational pursuit of self-interest by each player, is (R,R) or (D,D), regardless of the other player's choice, defection yields a higher payoff than cooperation, even though mutual cooperation would yield a better collective outcome. Therefore, participants face a trade-off between the best response and the norm-dependent utility.

Similarly, in the context of the Dictator Game, norm-dependent utility refers to the material gains adjusted by the disutility stemming from the deviation away from prevailing social norms. In the trust game, we can identify a Subgame Perfect Nash Equilibrium

(SPNE); the SPNE emerges as follows: the sender sends a portion of the initial amount to the receiver, demonstrating trust in the hopes of receiving a return. The receiver reciprocates by returning a proportion of the multiplied amount to the first player, honouring the trust placed in them. Suppose that the norm encourages the sender to send everything to the receiver. In the public goods game and modified public goods game, we observe a similar trade-off between maximising the collective outcome, self-interested outcomes and utility from complying with the norm.

The sender considers the best response of the receiver. Let

$$\frac{Du}{Da_i} = x_i^*(a_i) = max(V(\pi) + \gamma N + (\eta)(\pi_i(a_i) - \pi_j(a_j)) \tag{A.1}$$

Therefore, the sender also takes into account identity and decision of receiver.

If $x_i^*(a_i) \leq x_j^*(a_j)$ the sender maximises their utility.
If $x_i^*(a_i) > x_j^*(a_j)$ the sender chooses minimum amount deemed acceptable to the receiver.

Appendix A.2 - Instructions

Instructions for Games
Welcome!

The purpose of the research is to study how individuals make decisions in a number of economically relevant tasks and their beliefs about how others make decisions in those same tasks. You will earn a $2 participation fee (simply for participating). In addition, you may also earn extra money depending on your decisions. The task will take approximately 30 mins.

Decisions will be made using points. These points are converted into money. At the end of the study, the computer will randomly select one of your decisions for your payout A random draw will determine if your payout will depend on Decision A or B (decision as sender or receiver). This will decide your final payment. You will not know until the end of the study which task you will be paid for [1].

[1] Approximately two weeks later, we matched responses. Payments were transferred to bank accounts

You will not receive feedback on the outcome of any tasks. Please make your decisions in each task carefully. After the tasks are finished, you will also be asked to complete a questionnaire. The anonymity of your responses to all parts of the exercise (the tasks and questionnaire) are guaranteed.

Your decisions will be anonymous.

Thank you, and please continue.

Task 1.1

Here you are the SENDER.

In this scenario you will play with a blue and red card. How much money you will be paid depends on which card you play and which card the other player chooses. Please see the table below (5 is equal to 5 points).

Please click which card you will play if you play this game with a person from the following categories listed below[2].

Task 1.2

Which card do you think the person in the following categories will play, when playing

given by subjects. Participants were informed at the beginning of the session and in the flyers.

[2]Participants who choose an outcome with the least difference between themselves and the other player indicate a preference to cooperate (C, C); this is the Pareto-dominant strategy since there is no other option that gives both players a higher payoff. However, preferences to defect are indicated by participants choosing the outcome with the highest individual payoff, even if there is a significant difference between payoffs between participants (Cardenas & Carpenter 2008). Therefore, defecting is the dominant strategy (D, D)(Camerer 2011).

with you?

Task 2.1

Here you are the SENDER.

You will receive 10 points, you can decide how much of this to send to another player (the receiver.) The receiver can only accept the points you will send.

The points you will send will be the pay-out for the receiver.

The points you do not send, you will keep. Remember that the receiver is completely passive.

For each decision, you will have 10 points. What you do not send you keep. How much will you send to a person in the categories listed below?

Task 2.2

You are now the RECEIVER (completely passive).

For each decision, they (the sender) will have 10 points. What they do not send they keep.

How much would you expect to be sent from the individual senders listed below?

Task 3.1

Here you are the role of the SENDER.

You are given 10 points, you decide how much of this to send to the receiver (a person out of the categories below). The points you decide to give will be tripled (multiplied by 3) by us (the researchers).

Of this tripled amount the receiver can send back to you any points if they wish e.g. you send 4 points, it gets tripled by us to 12 points, the receiver can then send back up to 12 points if they wish.

For each decision, you will have 10 points. What you do not send you keep.

How much will you send to a person in the categories below?

Task 3.2

We now want you to act as the RECEIVER.

The senders listed below can send you any points between 0 and 10. The points sent to you will be tripled by us. You will receive the tripled points. Whatever points the sender does not send to you, the sender will keep for themselves.

How much would you expect from the individual senders listed below?

Task 3.3

You are now the receiver. How much would you send back to a [Ndebele/Shona/Zimbabwean] sender if you have received?

Task 4.1

Here you are the role of the SENDER.

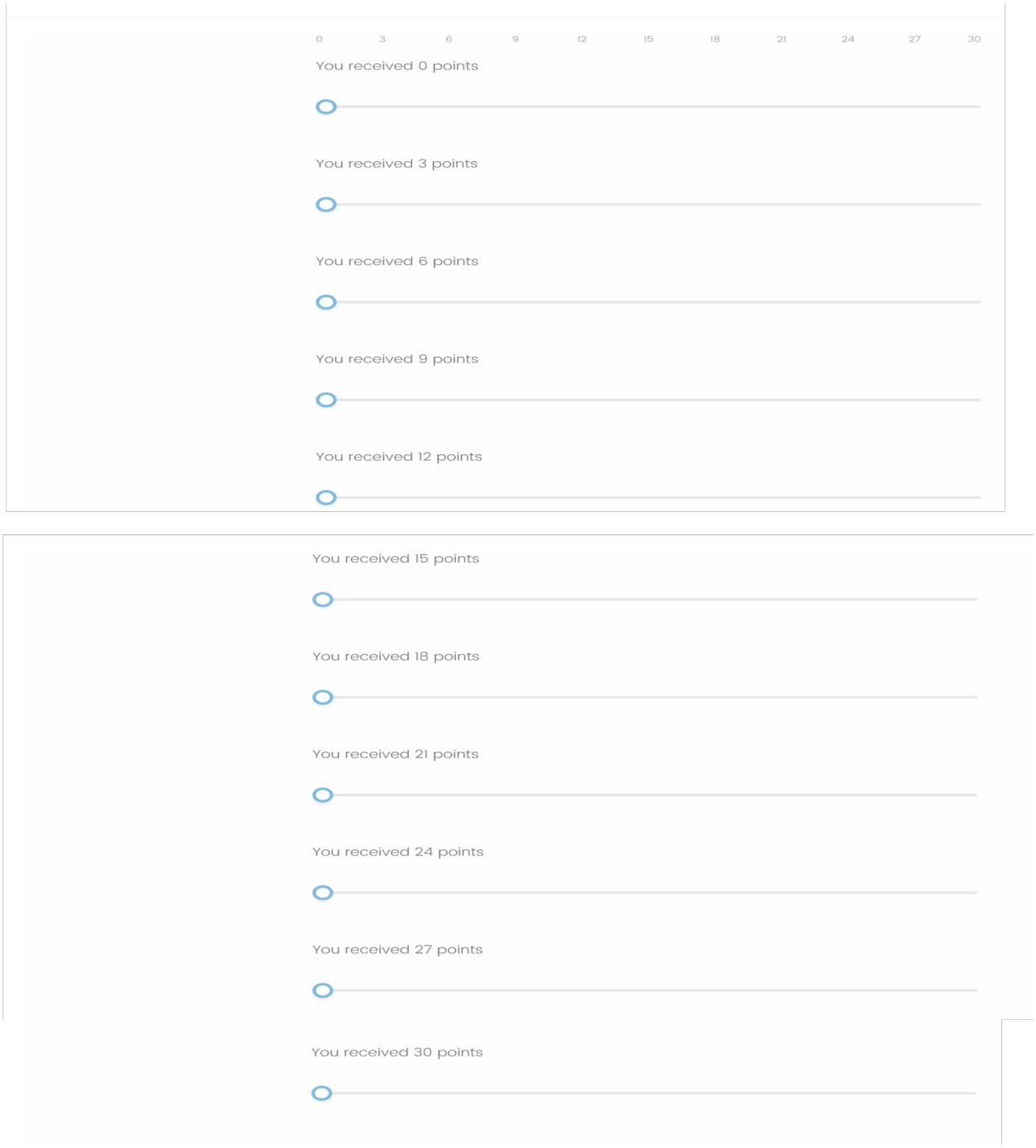


Figure A.1: Task 3.3. Responders decision in the Trust Game (in points)

Here you will play the game with 3 other group members. Each of you will have 10 points.The other participants know your identity.

Everybody can place up to 10 points on a virtual table. The money you do not put on the table, you will keep. You cannot see what other group members put on the table. The total points given by all four participants on the table will be multiplied by 1.5.

This multiplied points on the virtual table will be equally redistributed.

Your pay-out will be the money you will keep plus the equally re-distributed money from the table.

How many points will you put on the table when you play this game with three members in the categories below?

Task 4.2

You are with three other group members. The other participants know your identity.

For each decision, they will have 10 points. What they do not send they keep.

How many points do you think the member from the categories below will put on the table?

Task 4.3

You are now the receiver. You have 10 points.

You play the game with three other participants. Two of them are [Ndebele/Shona], the other participant is [Shona/Ndebele].

The total provided amount by all four participants will later be multiplied by 1.5.

How much would you contribute?

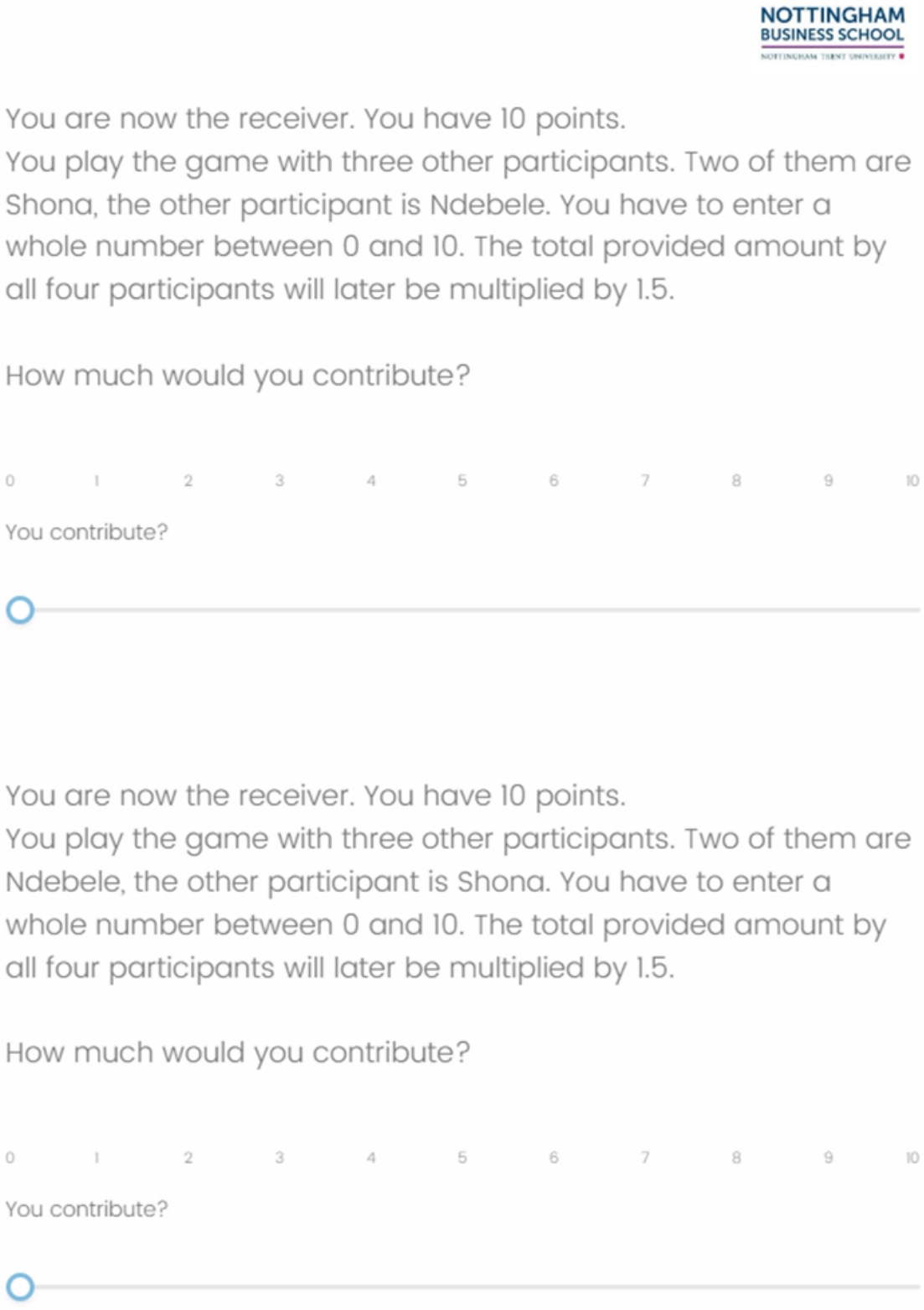


Figure A.2: Task 4.3. Responders decision in the Trust Game (in points)

Task Five

In this task, you will receive a description of a scenario.

You will be asked to consider each action. For each of the possible actions, you must decide whether taking that action would be "socially appropriate" or "socially inappropriate".

You will be given three options:

By socially appropriate, we mean behaviour that *you* think is the "correct" thing to do. By socially appropriate, we mean behaviour that *most people in your ethnicity* would agree is the "correct" thing to do By socially appropriate, we mean behaviour that *most people of the other ethnicity* would agree is the "correct" thing to do.

In each of your responses, we would like you to answer as truthfully as possible, based on your opinions of what is socially appropriate or socially inappropriate behaviour.

After this, your response is compared with the average response of what most Shona and Ndebele participants chose. If your evaluation is the same as this average, you will receive 5 points for this task; otherwise, you will receive zero.

We will go through an example situation to give you an idea of what the task will be like.

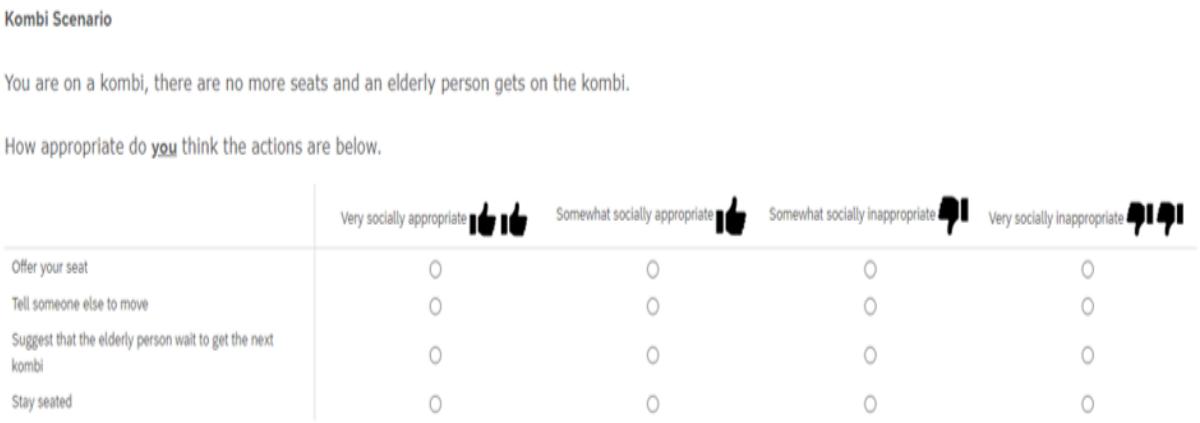


Figure A.3: Social norms elicitation task (Example scenario)

Appendix A.3 - Self-reported polarisation measure

To what extent do you align with the Shona identity?

- Very highly identify as
- Highly identify as
- Somewhat identify as
- No identification

To what extent do you align with the Ndebele identity?

- Very highly identify as
- Highly identify as
- Somewhat identify as
- No identification

Please say which of these words describe the way you think of yourself?

- Zimbabwean
- Shona
- Ndebele
- African

How proud are you to be [Ndebele/Shona]?

- Very proud
- Fairly proud
- Not very proud
- Not proud at all

Appendix A.4 - Additional analyses

Figure A.4 - Figure A.6 shows histograms of mean amounts sent by each identity to complement the main body[3]. In general, the patterns of offerings show a binomial distribution centred about 3 and 5.

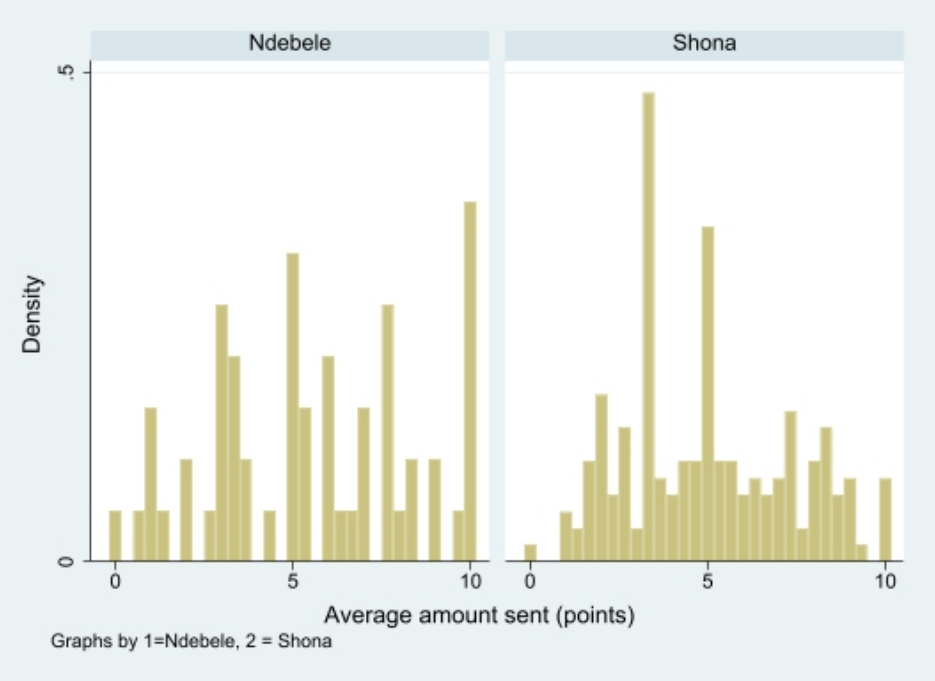


Figure A.4: Histograms of the mean amount sent by identity (DG)

[3]We exclude the Prisoner's Dilemma because of the binary decision strategy, see main text

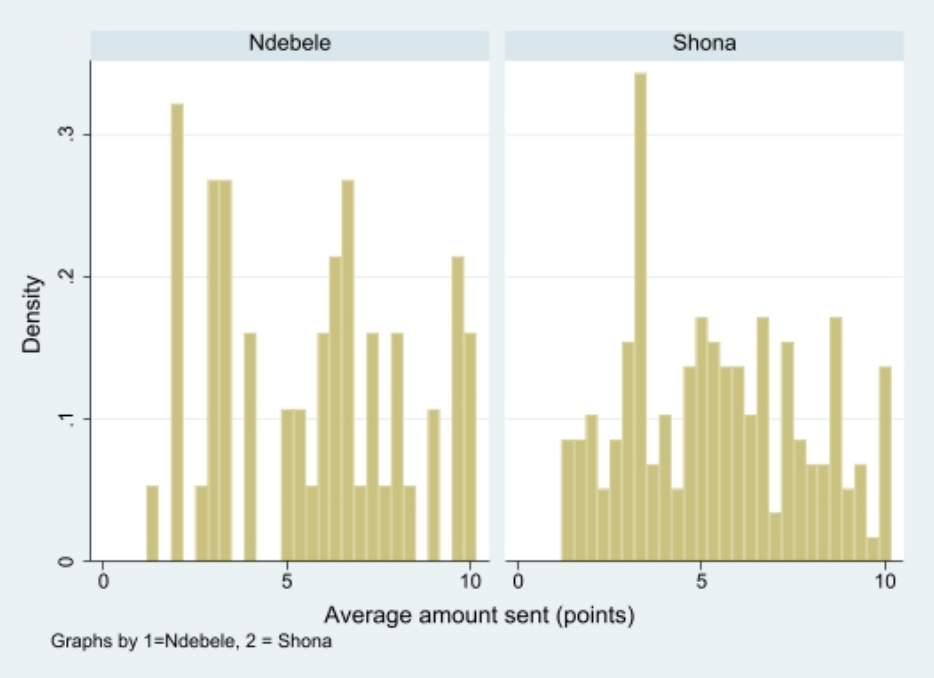


Figure A.5: Histograms of the mean amount sent by identity (TG)

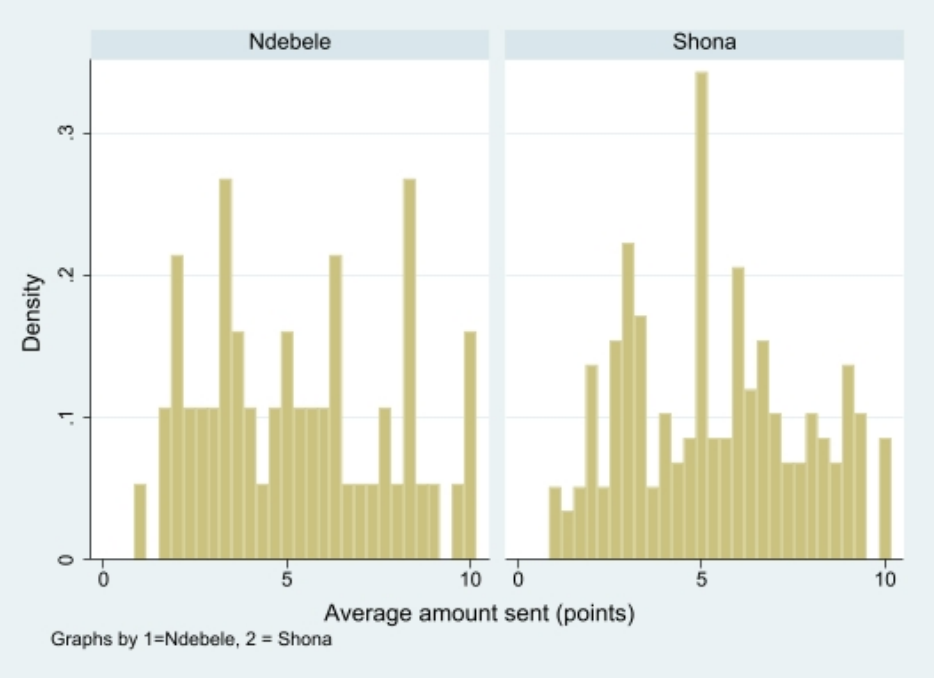


Figure A.6: Histograms of the mean amount sent by identity (PGG)

A.5.1 Alternative regressions of discriminatory behaviour

We also use a Tobit analysis as the amounts sent were constrained by zero and ten. Tobit model accounts for censoring, consistent with the literature we censor at both ends of the contributions. We include demographic controls; age, gender, risk preferences, religion, employment status, ethnicity, educational attainment and marital status. The results remain relatively unchanged from the main analysis.

Table A1: Tobit regression Dictator Game by receiver

	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(3)	(4)	(5)	(6)
Shona(=1, Ndebele=0)	-2.237***	-0.755*	0.232	0.628*	-0.818	0.509**
	(0.424)	(0.402)	(0.450)	(0.372)	(0.503)	(0.206)
Beliefs		0.527***		0.638***		0.765***
		(0.037)		(0.040)		(0.064)
Self-identity (Base=African)						
Ndebele		0.185		0.016		-0.929**
		(0.724)		(0.617)		(0.423)
Shona		0.187		-0.188		-0.330
		(0.575)		(0.689)		(0.354)
Zimbabwean		0.341		-0.135		-0.402
		(0.359)		(0.233)		(0.362)
Polarisation						
Shona align		-0.048		-0.372		-0.321
		(0.209)		(0.295)		(0.269)
Ndebele align		0.180		-0.056		0.096
		(0.143)		(0.192)		(0.143)
Proud		-0.018		0.608**		-0.051
		(0.189)		(0.261)		(0.295)
Controls	NO	YES	NO	YES	NO	YES
Constant	8.491***	2.260*	4.674***	-0.440	7.331***	-0.035

Table A1 – continued from previous page

	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(3)	(4)	(5)	(6)

	(0.529)	(1.197)	(0.675)	(0.932)	(0.777)	(1.428)
N	345	303	345	303	345	303

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table A2: Tobit regression Trust Game by receiver

	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(3)	(4)	(5)	(6)
Shona(=1, Ndebele=0)	-1.119**	-0.063	0.896***	0.494	-0.518	-0.619***
	(0.524)	(0.384)	(0.311)	(0.456)	(0.401)	(0.222)
Beliefs		0.665***		0.745***		0.837***
		(0.042)		(0.025)		(0.064)
Self-identity (Base=African)						
Ndebele		0.605		0.336		-0.332
		(0.444)		(0.335)		(0.413)
Shona		-0.895**		-0.521		0.344
		(0.383)		(0.452)		(0.367)
Zimbabwean		-0.066		0.144		-0.164
		(0.205)		(0.324)		(0.349)
Polarisation						
Shona align		0.260		-0.002		0.018
		(0.307)		(0.152)		(0.165)
Ndebele align		0.270		-0.205		-0.174
		(0.202)		(0.145)		(0.123)
Proud		-0.221		0.428**		0.148
		(0.223)		(0.168)		(0.100)
Controls	NO	YES	NO	YES	NO	YES

Table A2 – continued from previous page

	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(3)	(4)	(5)	(6)

Constant	7.184***	1.734**	4.008***	-0.156	7.349***	3.787***
	(0.628)	(0.881)	(0.428)	(0.964)	(0.461)	(0.793)
N	344	303	344	303	344	303

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table A3: Tobit regression Public Goods Game by receiver

	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(3)	(4)	(5)	(6)
Shona (=1, Ndebele=0)	-1.104**	-0.273	0.741	-0.016	0.106	0.094
	(0.507)	(0.268)	(0.483)	(0.519)	(0.636)	(0.333)
Beliefs		0.663***		0.725***		0.803***
		(0.050)		(0.039)		(0.034)
Self-identity (Base=African)						
Ndebele		-0.381		-1.073**		-1.142*
		(0.562)		(0.508)		(0.651)
Shona		-0.416		0.580		0.741
		(0.318)		(0.678)		(0.897)
Zimbabwean		-0.499***		-0.453**		0.035
		(0.133)		(0.177)		(0.166)
Polarisation						
Shona align		-0.271		-0.330		-0.267
		(0.202)		(0.217)		(0.203)
Ndebele align		-0.099		-0.229		-0.257
		(0.119)		(0.179)		(0.206)
Proud		-0.111		0.364***		0.027
		(0.147)		(0.135)		(0.122)

Table A3 – continued from previous page

	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(3)	(4)	(5)	(6)

Controls	NO	YES	NO	YES	NO	YES
Constant	6.980***	2.950***	4.220***	2.583*	5.862***	2.248**
	(0.632)	(0.699)	(0.705)	(1.440)	(0.913)	(1.015)
N	344	303	344	303	344	303

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

A.5.2 Additional analysis of treatment effects

The following figures present the transferred amount in each experimental game. The figures are counterparts to the figures shown in the main text for participants who are randomly assigned to the *Control*, *Heal* or *Divide* treatments. We include demographic controls, the control variables include age, gender, risk preferences, religion, employment status, ethnicity, educational attainment and marital status.

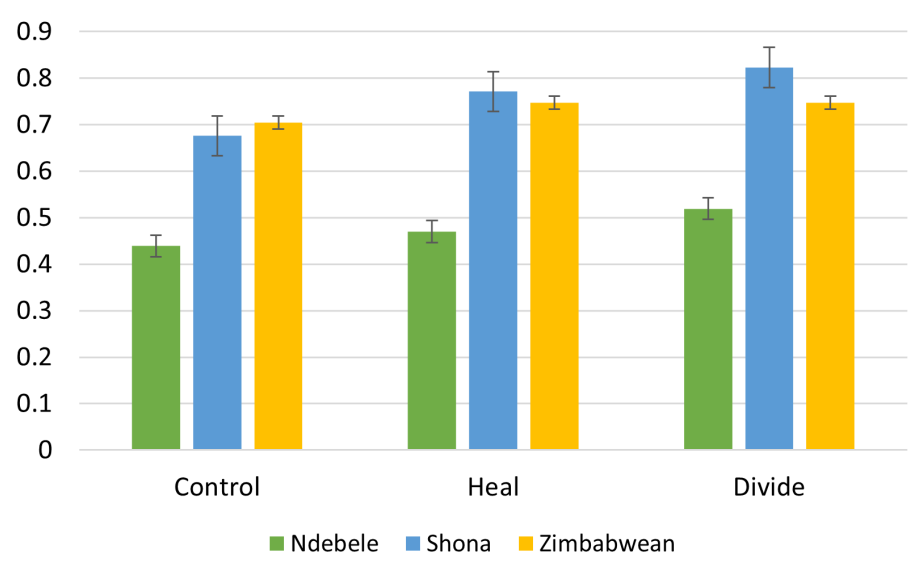


Figure A.7: Coordination in the Prisoner's Dilemma by treatment

Note: The figure shows the percentage of participants choosing coordinate (=1). The bars present coordination for each receiver (Ndebele, Shona and Zimbabwean, respectively). The first three columns are for the control group. The following three columns are for participants in the Heal treatment, and the final three columns are for participants in the Divide treatment. Error bars indicate 95% confidence intervals.

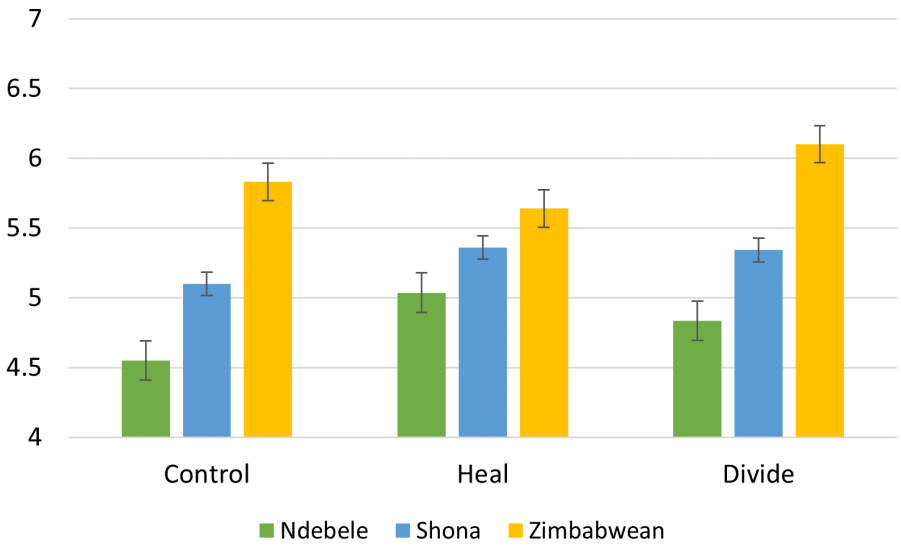


Figure A.8: Amounts sent in the Dictator game by treatment

Note: The figure shows the mean amount sent. The bars present cooperation for each receiver (Ndebele, Shona and Zimbabwean, respectively). The first three columns are for the control group. The following three columns are for participants in the Heal treatment, and the final three are for the Divide treatment. Error bars indicate 95% confidence intervals.

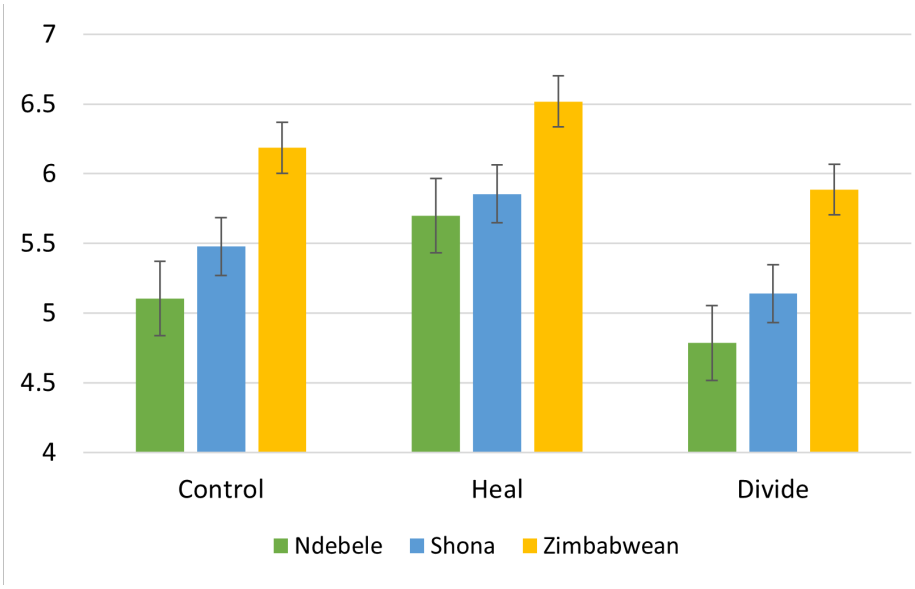


Figure A.9: Amount sent in the Trust Game by treatment

Note: The figure shows the mean amount sent. The bars present cooperation for each receiver (Ndebele, Shona and Zimbabwean, respectively). The first three columns are for the control group. The following three columns are for participants in the Heal treatment, and the final three are for the Divide treatment. Error bars indicate 95% confidence intervals.

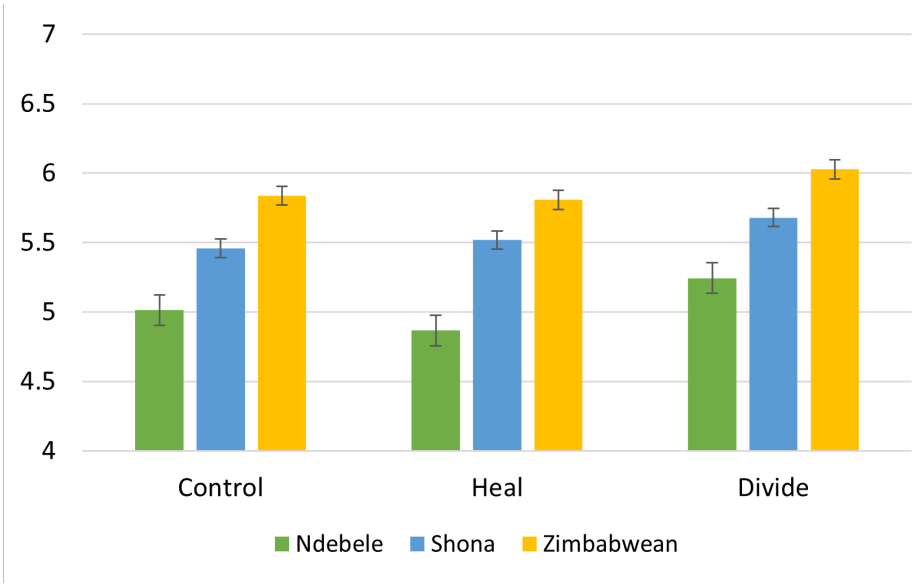


Figure A.10: Amount sent in the Public goods game by treatment (PGG)

Note: The figure shows the mean amount sent. The bars present cooperation for each receiver (Ndebele, Shona and Zimbabwean, respectively). The first three columns are for the control group. The following three columns are for participants in the Heal treatment, and the final three are for the Divide treatment. Error bars indicate 95% confidence intervals.

The following tables reports regression analysis of identity between the treatments, see Table D1 - Table D4. Column (1) presents results when the participant is Ndebele. Column (2) presents results when the participant is Shona. Our Probit regression models reveal that in the prisoner's dilemma, the *Heal* treatment increases coordination with the Shona receiver for both Ndebele and Shona participants. However, it reduces coordination with Ndebele receivers when the participant is Shona and reduces coordination with Zimbabwean receivers when the participant is Ndebele. The *Divide* treatment increases coordination with in-group members between Ndebele senders and Ndebele receivers. In addition, the Divide treatment increases coordination with Zimbabwean receivers from Shona senders. OLS regression in the dictator game shows that both treatments increase cooperation. The Heal treatment increases in-group cooperation. The Divide treatment increases in-group cooperation between Ndebele participants and between Ndebele senders and Zimbabwean receivers. One result of interest in the trust game is that both treatments increase inter-group trust; when the participant is Ndebele, they are more trusting toward Shona receivers. Similarly, with Ndebele participants to Zimbabwean receivers. Notably, in the public goods game, the Heal treatment backfires and reduces cooperation with Zimbabwean receivers from Ndebele senders.

Table D1: Probit Regression by treatment Prisoner's Dilemma

	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(1)	(2)	(1)	(2)
Heal	-0.154	-0.181*	0.582*	0.265***	-0.216**	0.051
	(0.147)	(0.086)	(0.259)	(0.059)	(0.078)	(0.074)
Divide	0.290*	-0.018	0.062	0.079	-0.131	0.197**
	(0.133)	(0.099)	(0.221)	(0.058)	(0.073)	(0.067)
Controls	YES	YES	YES	YES	YES	YES
Constant	-1.285	0.025	0.773	0.162	-0.836	0.353
	(0.889)	(0.339)	(1.227)	(0.341)	(0.998)	(0.343)
N	64	202	64	202	64	202

Note: Standard errors clustered at the subject level in parentheses.*** p<0.01,** p<0.05,* p<0.10.

Table D2: OLS Regression by treatment Dictator Game

	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(1)	(2)	(1)	(2)
Heal	1.724***	0.408	0.800	0.544*	0.492	0.084
	(0.481)	(0.352)	(0.425)	(0.251)	(0.505)	(0.232)
Divide	1.819*	0.213	0.280	0.469	1.993*	-0.074
	(0.916)	(0.290)	(0.948)	(0.434)	(0.981)	(0.474)
Controls	YES	YES	YES	YES	YES	YES
Constant	-0.040	1.924**	-1.903	2.324**	4.508	2.489
	(3.750)	(0.787)	(5.501)	(0.764)	(4.902)	(1.488)
N	65	238	65	238	65	238

Note: Standard errors clustered at the subject level in parentheses.*** p<0.01,** p<0.05,* p<0.10.

Table D3: OLS Regression by treatment Trust Game

	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(1)	(2)	(1)	(2)
Heal	1.801***	-0.050	1.611***	-0.625	0.860***	-0.487
	(0.444)	(0.214)	(0.393)	(0.361)	(0.166)	(0.509)
Divide	2.873***	0.471	3.020***	-0.104	2.645***	-0.247
	(0.807)	(0.321)	(0.485)	(0.238)	(0.414)	(0.152)
Controls	YES	YES	YES	YES	YES	YES
Constant	3.502	1.191	2.969	2.587**	4.184***	2.055***
	(3.015)	(0.920)	(2.005)	(1.041)	(0.678)	(0.592)
N	65	238	65	238	65	238

Note: Standard errors clustered at the subject level in parentheses.*** p<0.01,** p<0.05,* p<0.10.

Table D4: OLS Regression by treatment Public Goods Game

	Transfer to Ndebele		Transfer to Shona		Transfer to Zim	
	(1)	(2)	(1)	(2)	(1)	(2)
Heal	-0.233	-0.058	-0.091	-0.357	-1.284**	-0.252
	(0.543)	(0.329)	(0.934)	(0.579)	(0.448)	(0.551)
Divide	-0.474	0.299	-0.627	-0.086	0.091	-0.255
	(0.678)	(0.375)	(0.811)	(0.385)	(0.478)	(0.545)
Controls	YES	YES	YES	YES	YES	YES
Constant	3.601	2.762***	1.302	4.763***	3.001	2.735**
	(4.122)	(0.472)	(2.138)	(0.686)	(2.007)	(0.918)
N	65	238	65	238	65	238

Note: Standard errors clustered at the subject level in parentheses.*** p<0.01,** p<0.05,* p<0.10.

A.5.3 Additional analysis of norm effects

To identify the impact of individual norm evaluations, we replicate regressions in the main text but include each evaluation the evaluator considers on mean cooperativeness across each game, see Table D5. The dependent variable is the pooled mean contributions for all receivers (Ndebele, Shona and Zimbabwean) in each game. The (Nd) refers to evaluations of the Ndebele majority. The (Sh) refers to evaluations of the Shona majority. The (Own) refers to an individual's evaluation. We include demographic controls. The control variables include age, gender, risk preferences, religion, employment status, ethnicity, educational attainment and marital status.

Table D5: OLS Regression on effects of social norms on mean contributions in games

	PD	DG	TG	PGG
	(1)	(2)	(3)	(4)
Shona(=1, Ndebele=0)	-0.059	0.130	-0.358	-0.237
	(0.053)	(0.203)	(0.389)	(0.439)
normfamily(own)	-0.009	0.642*	-0.168	-0.036
	(0.028)	(0.291)	(0.223)	(0.150)
normfamily(Sh)	0.039	-0.328*	0.330	-0.032
	(0.045)	(0.152)	(0.179)	(0.145)
normfamily(Nd)	0.003	-0.264*	-0.168	0.190
	(0.024)	(0.124)	(0.181)	(0.108)
normpay(own)	0.003	-0.072	-0.127	0.060
	(0.027)	(0.363)	(0.329)	(0.243)
normpay(Sh)	0.044**	0.157	-0.012	-0.260
	(0.015)	(0.100)	(0.284)	(0.214)
normpay(Nd)	-0.025	-0.155	-0.050	0.002
	(0.022)	(0.181)	(0.308)	(0.233)
normtax(own)	-0.010	0.310	-0.112	-0.022
	(0.029)	(0.282)	(0.134)	(0.304)
normtax(Sh)	-0.032	0.172	-0.473**	0.111
	(0.025)	(0.331)	(0.193)	(0.400)
normtax(Nd)	0.026	-0.601**	0.210	-0.059
	(0.045)	(0.183)	(0.177)	(0.235)
normbribe1(own)	0.029	-0.191	-0.008	-0.137
	(0.029)	(0.257)	(0.134)	(0.160)
normbribe1(Sh)	-0.068**	-0.026	0.323**	0.209
	(0.023)	(0.177)	(0.115)	(0.152)
normbribe1(Nd)	0.021	0.518**	-0.106	0.424*
	(0.024)	(0.192)	(0.123)	(0.185)
normbribe2(own)	0.064*	0.137	-0.288	-0.462***
	(0.031)	(0.564)	(0.162)	(0.084)
normbribe2(Sh)	-0.044	-0.360	-0.723**	-0.248
	(0.031)	(0.212)	(0.205)	(0.269)

Continued on next page

Table D5 – continued from previous page

	PD	DG	TG	PGG
	(1)	(2)	(3)	(4)
normbribe2(Nd)	0.027	-0.081	0.231	0.398**
	(0.021)	(0.091)	(0.225)	(0.137)
Heal	0.162***	-0.088	-1.839***	-0.306
	(0.035)	(0.341)	(0.417)	(0.195)
Divide	0.253***	1.057***	1.048**	-1.959***
	(0.069)	(0.152)	(0.419)	(0.482)
Controls	YES	YES	YES	YES
Constant	0.572***	-0.525	0.706	3.037*
	(0.145)	(1.166)	(1.570)	(1.388)
N	263	189	189	189

Note: Standard errors clustered at the subject level.*** p<0.01,** p<0.05,* p<0.10.

B

Appendix B for Chapter 3

Appendix B.1 - Instructions

Instructions for Games

Welcome!

The purpose of the research is to study how individuals make decisions in a number of economically relevant tasks and their beliefs about how others make decisions in those same tasks. You will be presented with several economic scenarios in which you are tasked to make a value judgement and a decision on the social appropriateness of different scenarios. As you do this, you will receive points, depending on the response you have provided to the survey.

You will earn a £2.50 participation fee (simply for participating). In addition, you may also earn extra money depending on your decisions and the choices of others in the study. The task will take approximately 30 mins.

At the end of the study, the computer will match your responses with another participants answers and randomly select one of your decisions for your payout. This will decide your final payment [1].

You will not know until the end of the study which task you will be paid for; please carefully make your decisions in each task. After the tasks are finished, you will also be asked to complete a questionnaire. The anonymity of your responses to all parts of the exercise (the tasks and questionnaire) is guaranteed.

Your decision will be anonymous

Thank you, and please continue.

Task 1.1

Here you are the SENDER.

In this scenario you will play with a blue and red card. How much money you will be paid depends on which card you play and which card the other player chooses. Please see the table below (5 is equal to 5 points).

Please click which card you will play if you play this game with a person from the following categories listed below[2].

[1] Approximately two days later, we matched responses. Payments were transferred to bank accounts via Prolific Academic given to by subjects. Participants were informed that their responses were matched with the responses of another session. We collected the receivers' responses for all games. The collected responders' choices were matched with the senders' responses to determine the payoff. No additional responses were collected for the Prisoners' Dilemma, and the matching took place with the responses from a different session. Participants were informed at the beginning of the session and in the flyers.

[2] Participants who choose an outcome with the least difference between themselves and the other player indicate a preference to cooperate (C, C); this is the Pareto-dominant strategy since there is no other option that gives both players a higher payoff. However, preferences to defect are indicated by participants choosing the outcome with the highest individual payoff, even if there is a significant difference between payoffs between participants (Cardenas & Carpenter 2008). Therefore, defecting is the dominant strategy (D, D)(Camerer 2011).

Task 1.2

Which card do you think the person in the following categories will play, when playing with you?

Task 2.1

Here you are the SENDER.

You will receive 10 points, you can decide how much of this to send to another player (the receiver.) The receiver can only accept the points you will send.

The points you will send will be the pay-out for the receiver.

The points you do not send, you will keep. Remember that the receiver is completely passive.

For each decision, you will have 10 points. What you do not send you keep. How much will you send to a person in the categories listed below?

Task 2.2

You are now the RECEIVER (completely passive).

For each decision, they (the sender) will have 10 points. What they do not send they keep.

How much would you expect to be sent from the individual senders listed below?

Task 3.1

Here you are the role of the SENDER.

You are given 10 points, you decide how much of this to send to the receiver (a person out of the categories below). The points you decide to give will be tripled (multiplied by 3) by us (the researchers).

Of this tripled amount the receiver can send back to you any points if they wish e.g. you send 4 points, it gets tripled by us to 12 points, the receiver can then send back up to 12 points if they wish.

For each decision, you will have 10 points. What you do not send you keep.

How much will you send to a person in the categories below?

Task 3.2

We now want you to act as the RECEIVER.

The senders listed below can send you any points between 0 and 10. The points sent to you will be tripled by us. You will receive the tripled points. Whatever points the sender does not send to you, the sender will keep for themselves.

How much would you expect from the individual senders listed below?

Task 3.3

You are now the receiver. How much would you send back to a [Leaver/Remainer/UK citizen] sender if you have received?

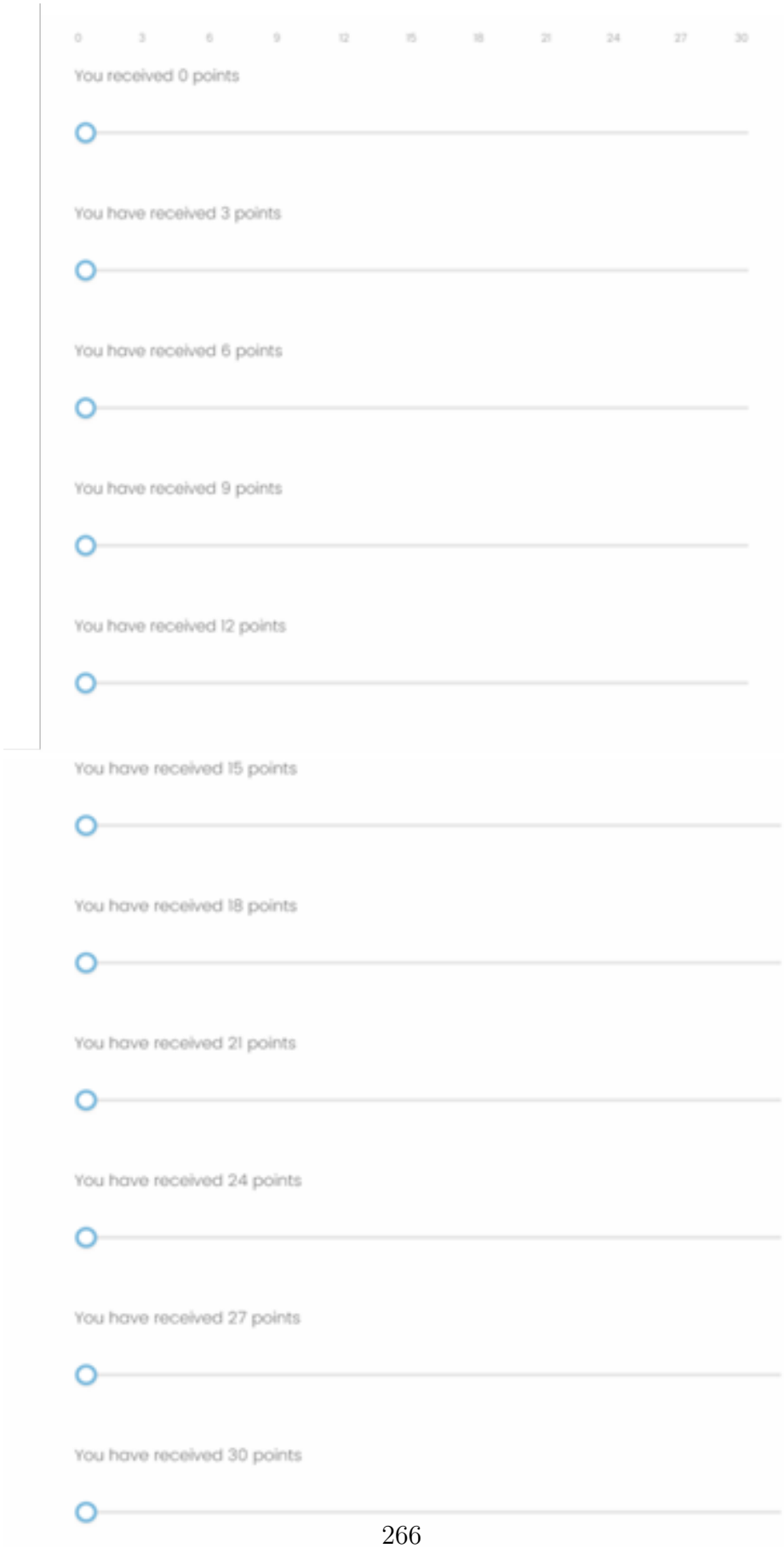


Figure B.1: Trust Game responder decision

Task 4.1

Here you are the role of the SENDER.

Here you will play the game with 3 other group members. Each of you will have 10 points.The other participants know your identity.

Everybody can place up to 10 points on a virtual table. The money you do not put on the table, you will keep. You cannot see what other group members put on the table. The total points given by all four participants on the table will be multiplied by 1.5.

This multiplied points on the virtual table will be equally redistributed.

Your pay-out will be the money you will keep plus the equally re-distributed money from the table.

How many points will you put on the table when you play this game with three members in the categories below?

Task 4.2

You are with three other group members. The other participants know your identity.

For each decision, they will have 10 points. What they do not send they keep.

How many points do you think the member from the categories below will put on the table?

Task 4.3

You are now the receiver. You have 10 points.

You play the game with three other participants. Two of them are [Leaver/Remainer], the other participant is [Leaver/Remainer].

The total provided amount by all four participants will later be multiplied by 1.5.

How much would you contribute?

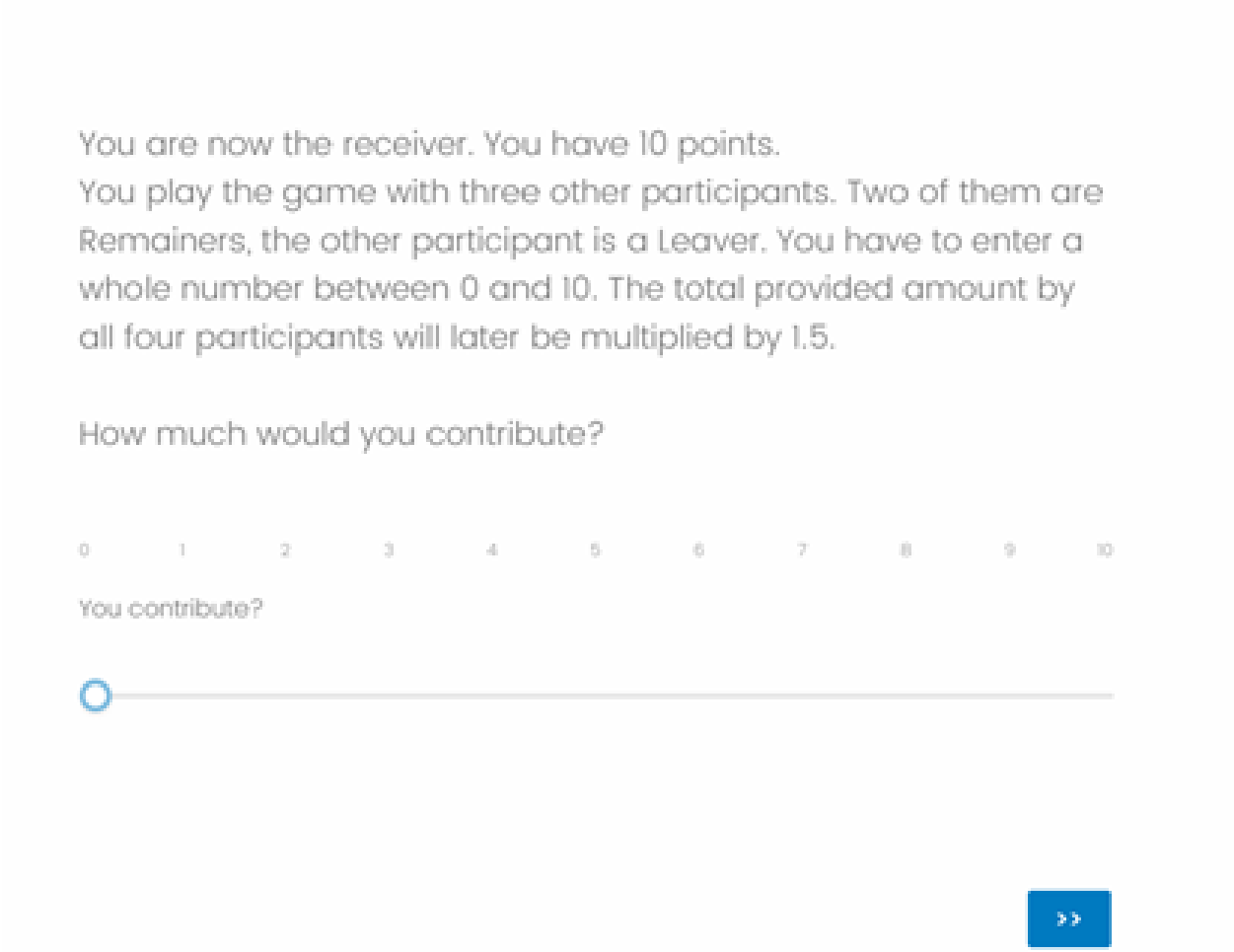


Figure B.2: Public Goods Game responder decision

Task 5.1

New Task

Here you are the SENDER.

You will interact with one other group member. The other member will be from one of the categories below. Both of you will have 10 points. The other participant will know who they are interacting with, as either a Leaver, Remainer, Non-voter, UK citizen, Scottish, English, Welsh, and Northern Irish person.

You can both place up to 10 points on a virtual table. The points you do not put on the table, you will keep. You cannot see what the other group member puts on the table. The total points given by the two group members on the table will be multiplied by 1.5.

E.g. If there are 0 points on the table, then 0 points multiplied by 1.5 equals 0 points. If there are 20 points on the table, then 20 points multiplied by 1.5 equals 30 points.

The multiplied points on the virtual table will be equally redistributed to both members.

The points you will earn will be the points you chose to keep plus the equally redistributed points from the table. For each decision, you will have 10 points. What you do not send you keep.

How many points will you put on the table when interacting with one person from the categories below?

Task 5.2

You are now the RECEIVER.

How many points do you think the member from the categories below will put on the table when interacting with you?

Task Six

In this task, you will receive a description of a scenario.

You will be asked to consider each action. Please decide, for each of the possible actions, whether taking that action would be "socially appropriate" or "socially inappropriate".

Please consider your own thoughts in each scenario, as well as what you think the actions of Leavers and Remainers is.

In other words, you will be asked to consider each scenario three times:

By socially appropriate, we mean actions that you / Leavers / Remainers think is the "correct" thing to do.

By socially inappropriate, we mean actions that you / Leavers / Remainers would agree is "not the correct" thing to do.

We would like you to answer as truthfully as possible in each of your responses.

After this, your response is compared with the answer that most Leave, and Remain participants chose. If your evaluation is the same as the most common answer, you will receive 5 points for this task; otherwise, you will earn 0 points.

We will go through an example situation to give you an idea of what the task will be like.

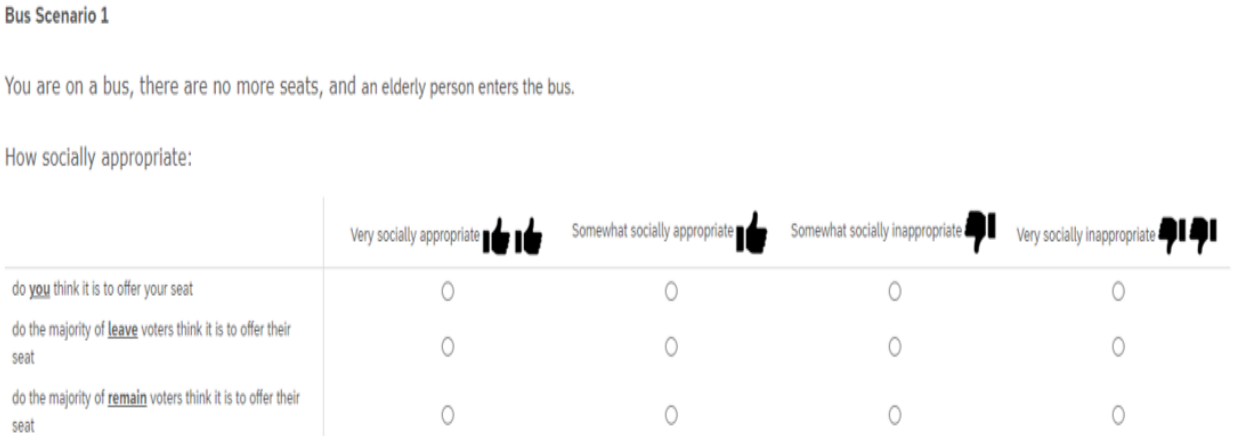


Figure B.3: Social norms elicitation task (Example scenario)

Appendix B.2 - Self-reported polarisation measures

Following on from the 2016 EU Referendum, do you feel like a winner or a loser?

- Winner
- Loser

Now that Britain has left the EU, would you vote the same?

- Yes
- No
- Unsure

To what extent do you align with the reported previous identity?

- Very highly identify as
- Highly identity as
- Somewhat identify as
- No identification

How proud are you to have vote reported previous voting decision?

- Very proud
- Fairly proud
- Not very proud
- Not proud at all

Appendix B.3 - Additional analyses

Figure B.4 - Figure B.7 shows histograms of mean amounts sent by identity to complement the main body[3]. In general, the patterns of offerings show a binomial distribution centred around 5 and 10.

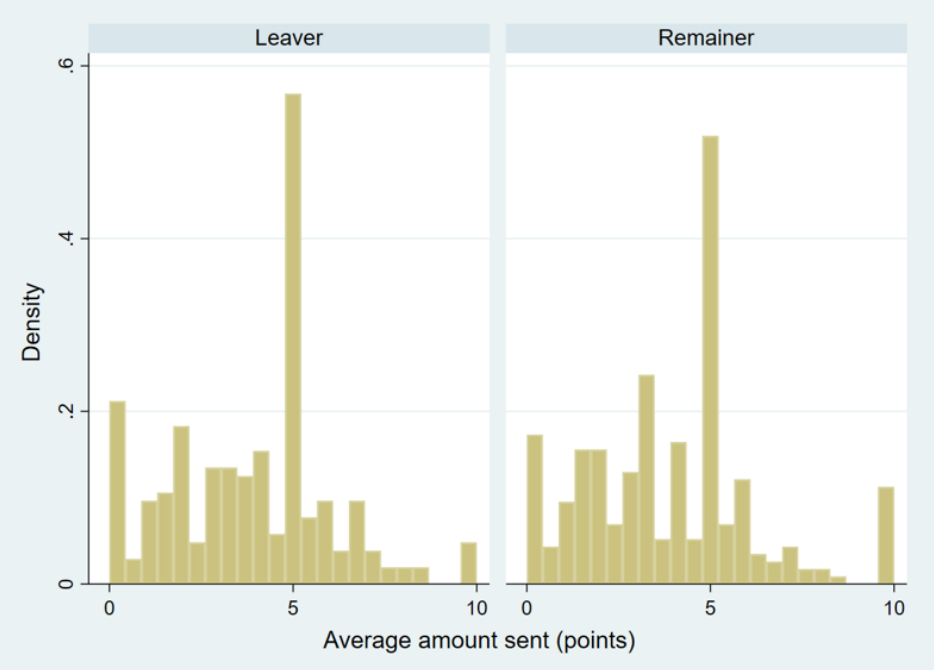


Figure B.4: Histogram of amounts sent in the Dictator Game

[3]We exclude the Prisoner's Dilemma because of the binary decision strategy, see main text.

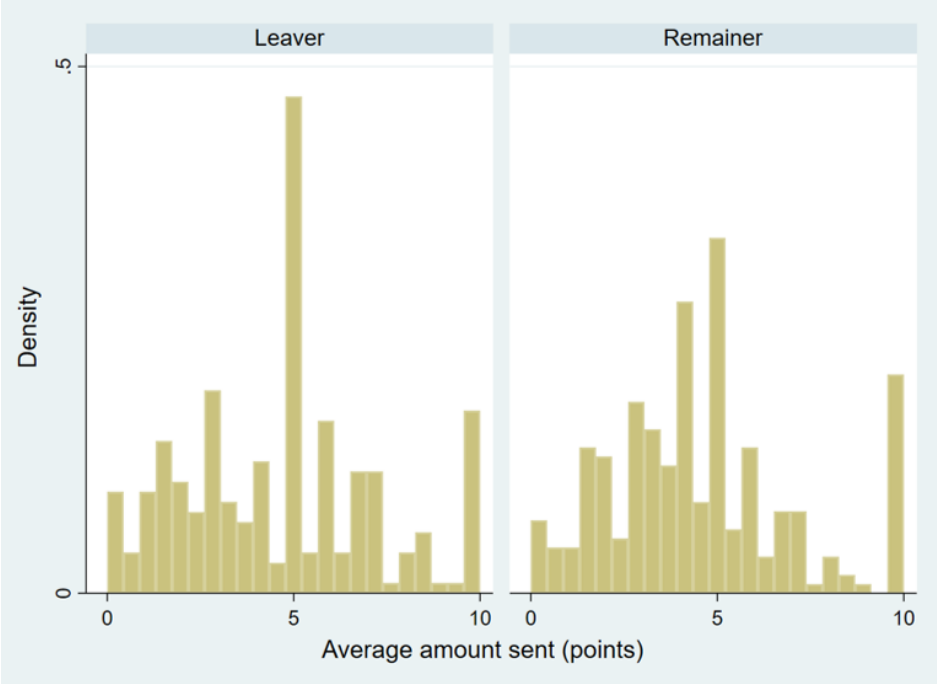


Figure B.5: Histogram of the amounts sent in the Trust Game

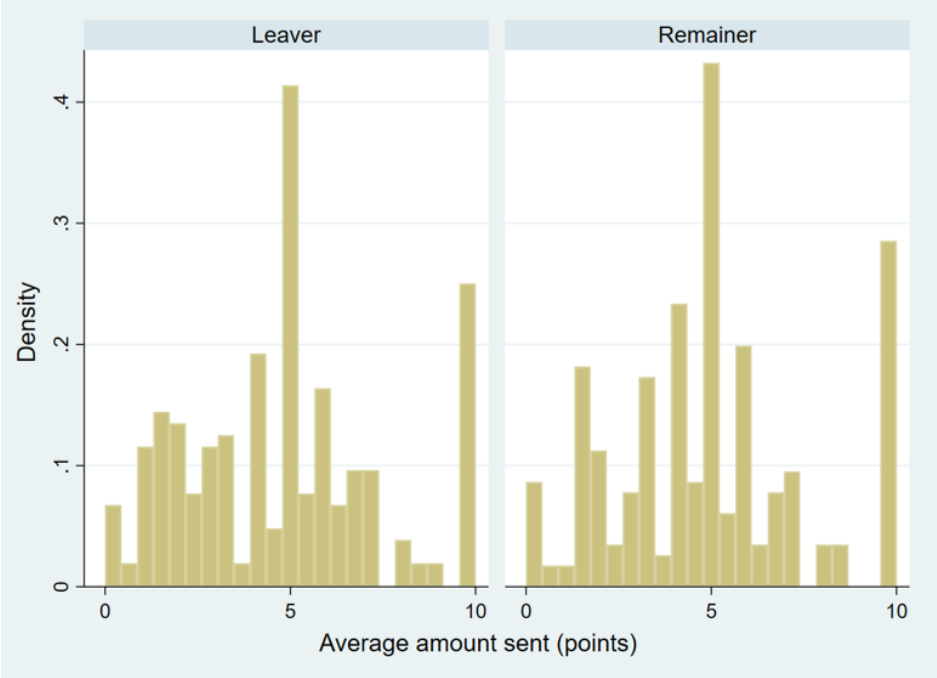


Figure B.6: Histogram of the amounts sent in the Public Goods Game

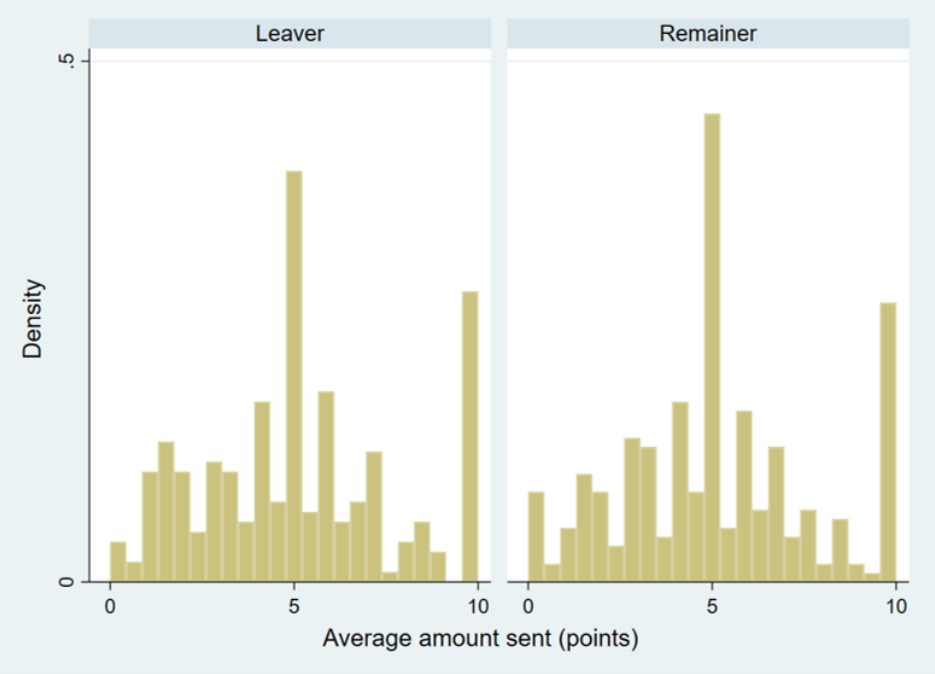


Figure B.7: Histogram of the amounts sent in the Modified Public Goods Game

B.3.1 Alternative regressions of discriminatory behaviour

We also use a Tobit regression model that accounts for censoring, consistent with the literature we censor at both ends of the contributions. We include demographic controls; age, gender, risk preferences, religion, employment status, ethnicity, educational attainment and marital status. The results are relatively unchanged from the main text.

Table B1: Tobit regression Dictator Game by contribution

	Transfer Leaver		Transfer Remainer		Transfer UK	
	(1)	(2)	(3)	(4)	(5)	(6)
Leaver(=1, Remainer=0)	0.357	0.090	-0.737***	0.146	-0.238	-0.161
	(0.260)	(0.333)	(0.243)	(0.324)	(0.186)	(0.282)
Beliefs	0.987***	0.970***	0.877***	0.888***	0.913***	0.922***
	(0.057)	(0.054)	(0.051)	(0.048)	(0.238)	(2.911)

Table B1 – continued from previous page

	Transfer Leaver		Transfer Remainer		Transfer UK	
	(1)	(2)	(3)	(4)	(5)	(6)

Polarisation						
Vote proud		-0.109		-0.019		-0.002
		(0.139)		(0.128)		(0.111)
Winner		-0.114		-1.016***		-0.440*
		(0.286)		(0.289)		(0.252)
Align		0.609**		-0.458**		-0.135
		(0.272)		(0.226)		(0.204)
Controls	NO	YES	NO	YES	NO	YES
Constant	-0.222	-5.337	1.043***	2.166	0.593**	-2.550
	(0.208)	(3.864)	(0.273)	(1.531)	(0.238)	(2.911)
Observations	505	505	505	505	505	505

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table B2: Tobit regression Trust Game by contribution

	Transfer Leaver		Transfer Remainer		Transfer UK	
	(1)	(2)	(3)	(4)	(5)	(6)
Leaver(=1, Remainer=0)	0.116	-0.366	-0.507*	0.030	-0.030	0.027
	(0.284)	(0.389)	(0.282)	(0.402)	(0.180)	(0.279)
Beliefs (Leaver)	1.128***	1.123***	0.947***	0.942***	1.043***	1.044***
	(0.051)	(0.050)	(0.052)	(0.051)	(0.035)	(0.036)
Polarisation						
Vote proud		-0.077		-0.230		-0.037
		(0.139)		(0.167)		(0.117)
Winner		0.599**		-0.870***		-0.314
		(0.304)		(0.319)		(0.241)
Align		0.195		-0.057		0.188

Table B2 – continued from previous page

	Transfer Leaver		Transfer Remainer		Transfer UK	
	(1)	(2)	(3)	(4)	(5)	(6)

		(0.266)		(0.271)		(0.205)
Controls	NO	YES	NO	YES	NO	YES
Constant	-0.301	0.443	0.934***	0.213	0.246	1.456
	(0.204)	(2.564)	(0.315)	(3.773)	(0.194)	(1.241)
Observations	505	505	505	505	505	505

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table B3: Tobit regression Public Goods Game by contribution

	Transfer Leaver		Transfer Remainer		Transfer UK	
	(1)	(2)	(3)	(4)	(5)	(6)
Leaver(=1, Remainer=0)	0.156	-0.226	-0.698**	-0.088	-0.130	-0.144
	(0.274)	(0.371)	(0.271)	(0.350)	(0.174)	(0.285)
Beliefs	1.163***	1.158***	1.013***	1.007***	1.034***	1.037***
	(0.050)	(0.050)	(0.049)	(0.050)	(0.036)	(0.036)
Polarisation						
Vote proud		-0.243*		-0.152		-0.096
		(0.142)		(0.144)		(0.103)
Winner		0.270		-0.982***		-0.006
		(0.339)		(0.326)		(0.288)
Align		0.502*		-0.134		0.034
		(0.287)		(0.235)		(0.198)
Controls	NO	YES	NO	YES	NO	YES
Constant	-0.180	1.401	0.808***	3.683*	0.386**	2.317**
	(0.217)	(2.222)	(0.285)	(2.083)	(0.185)	(1.029)
Observations	505	505	505	505	505	505

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table B4: Tobit regression Public Goods Game by contribution

	Transfer Leaver		Transfer Remainer		Transfer UK	
	(1)	(2)	(3)	(4)	(5)	(6)
Leaver(=1, Remainer=0)	0.271	-0.013	-0.433*	-0.014	-0.088	0.217
	(0.272)	(0.359)	(0.239)	(0.360)	(0.199)	(0.286)
Beliefs	1.112***	1.104***	1.064***	1.060***	1.070***	1.072***
	(0.052)	(0.051)	(0.044)	(0.045)	(0.040)	(0.039)
Polarisation Vote proud		-0.146		-0.132		-0.023
		(0.168)		(0.169)		(0.128)
Winner		0.212		-0.288		-0.487*
		(0.319)		(0.323)		(0.263)
Align		0.072		-0.392		-0.106
		(0.291)		(0.244)		(0.213)
Controls	NO	YES	NO	YES	NO	YES
Constant	0.186	0.951	0.677**	3.168	0.256	2.170
	(0.230)	(2.430)	(0.270)	(2.093)	(0.233)	(2.588)
Observations	505	505	505	505	505	505

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

B.3.2 Robustness checks for intervention

The following figures present the transferred amount in each experimental game. The figures are counterparts to the figures shown in the main text for participants who are randomly assigned to the Control, *Heal* or *Divide* treatments.

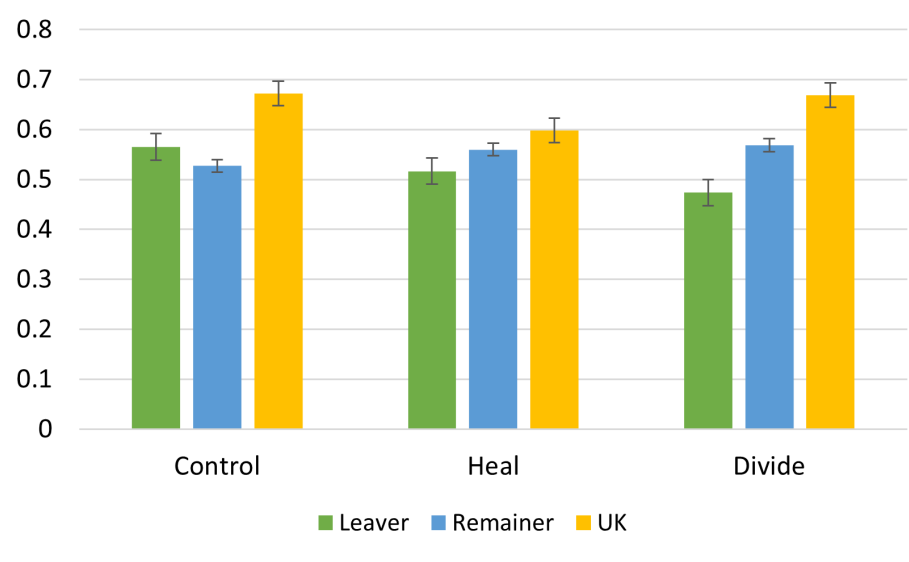


Figure B.8: Prisoner's Dilemma

Note: The figure shows the percentage of participants choosing coordinate (=1). The bars present coordination for each receiver (Leaver, Remainer, and the UK, respectively). The first three columns are for the control group. The following three columns are for participants in the Heal treatment, and the final three are for the Divide treatment. Error bars indicate 95% confidence intervals.

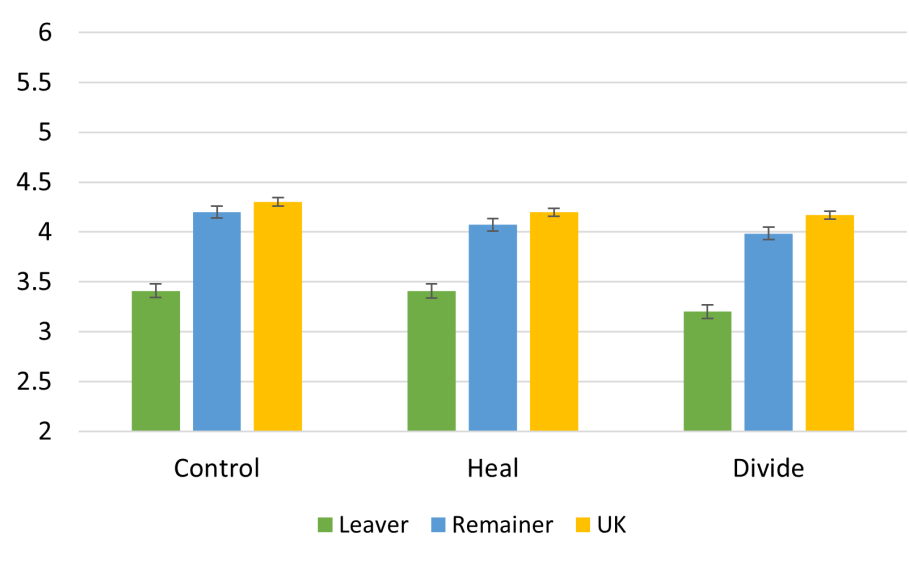


Figure B.9: Dictator Game

Note: The figure shows the mean amount sent. The bars present cooperation for each receiver (Leaver, Remainer, and the UK, respectively). The first three columns are for the control group. The following three columns are for participants in the Heal treatment, and the final three are for the Divide treatment. Error bars indicate 95% confidence intervals.

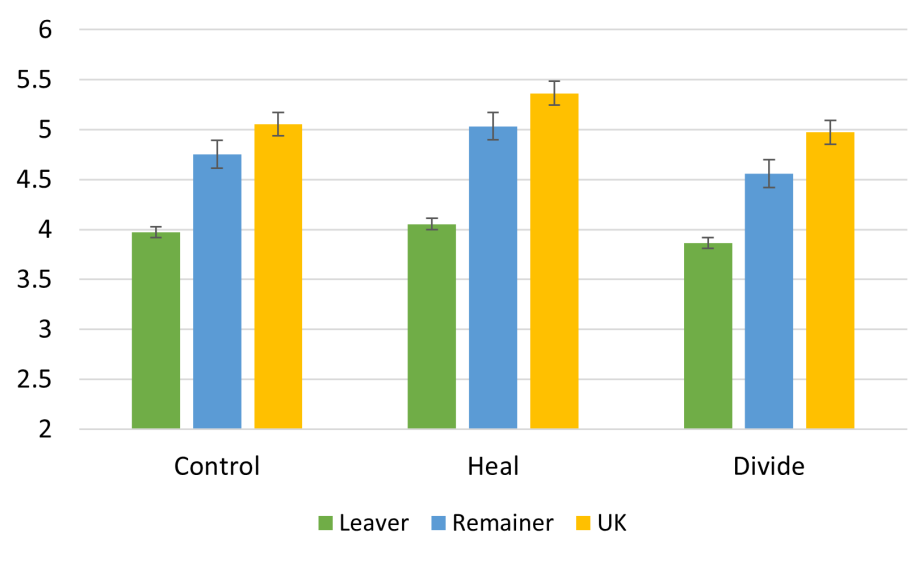


Figure B.10: amount sent in the Trust game by treatment

Note: The figure shows the mean amount sent. The bars present cooperation for each receiver (Leaver, Remainer, and the UK, respectively). The first three columns are for the control group. The following three columns are for participants in the Heal treatment, and the final three are for the Divide treatment. Error bars indicate 95% confidence intervals.

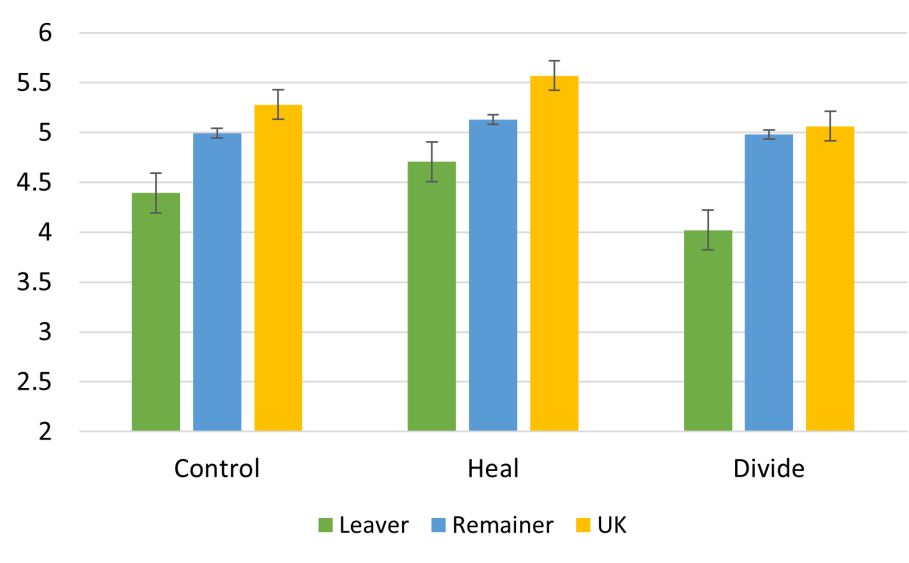


Figure B.11: amount sent in the Public Goods Game by treatment

Note: The figure shows the mean amount sent. The bars present cooperation for each receiver (Leaver, Remainer, and the UK, respectively). The first three columns are for the control group. The following three columns are for participants in the Heal treatment, and the final three are for participants in the Divide treatment. Error bars indicate 95% confidence intervals.

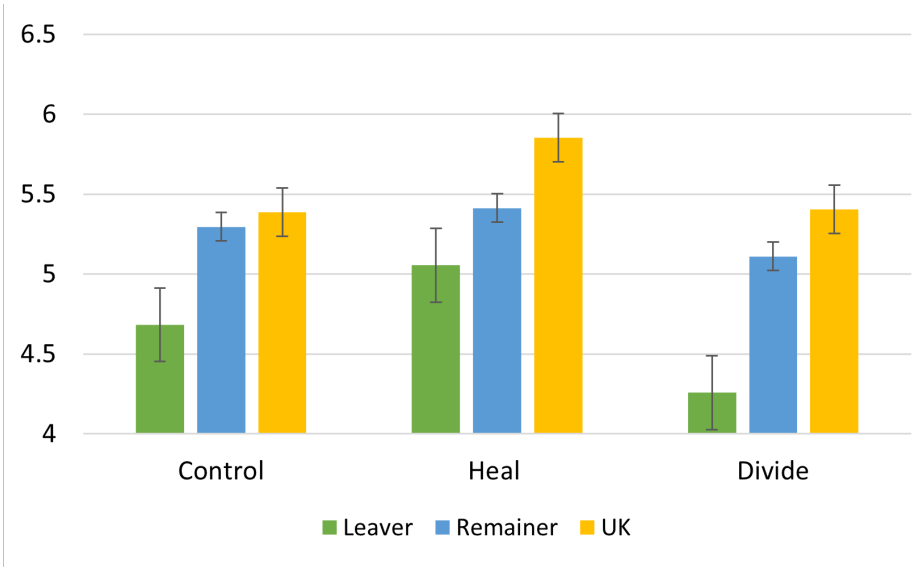


Figure B.12: Public Goods Game

Note: The figure shows the mean amount sent. The bars present cooperation for each receiver (Leaver, Remainer, and the UK, respectively). The first three columns are for the control group. The following three columns are for participants in the Heal treatment, and the final three are for participants in the Divide treatment. Error bars indicate 95% confidence intervals.

To further disentangle treatment effects, we complement regressions in the main body by separating the regressions by the participant's identity. Column (1) if Leaver. Column (2) if Remainer. The findings are similar to results in chapter 3. We include demographic controls; age, gender, risk preferences, religion, employment status, ethnicity, educational attainment and marital status.

Table B5: Probit regression Prisoner's dilemma by contribution and identity

	Transfer to Leaver		Transfer to Remainer		Transfer to UK	
	(L)	(R)	(L)	(R)	(L)	(R)
Heal	-0.287	-0.119	0.149	0.138	-0.529**	-0.123
	(0.228)	(0.211)	(0.219)	(0.225)	(0.242)	(0.220)
Divide	-0.725***	-0.054	0.231	0.212	-0.493**	0.241
	(0.220)	(0.200)	(0.217)	(0.207)	(0.238)	(0.216)
Controls	YES	YES	YES	YES	YES	YES

Table B5 – continued from previous page

	Transfer to Leaver		Transfer to Remainer		Transfer to UK	
	(L)	(R)	(L)	(R)	(L)	(R)

Constant	6.078***	5.671***	-4.513***	-7.062***	7.236***	-4.805***
	(0.907)	(1.226)	(0.855)	(1.275)	(1.026)	(1.190)
Observations	238	266	238	266	238	266

Note: Robust standard errors clustered by the session in parentheses.*** p<0.01,** p<0.05,* p<0.10

Table B6: OLS regression Dictator Game by contribution and identity

	Transfer to Leaver		Transfer to Remainer		Transfer to UK	
	(L)	(R)	(L)	(R)	(L)	(R)
Heal	0.197	-0.123	0.269	-0.191	0.017	-0.182
	(0.249)	(0.304)	(0.265)	(0.294)	(0.258)	(0.279)
Divide	-0.093	-0.097	-0.038	0.139	-0.028	0.135
	(0.307)	(0.263)	(0.293)	(0.231)	(0.280)	(0.237)
Controls	YES	YES	YES	YES	YES	YES
Constant	-2.891	-0.176	4.057***	-0.981	-2.030	-0.410
	(3.030)	(1.781)	(1.501)	(1.388)	(3.110)	(1.497)
Observations	239	266	239	266	239	266

Note: Robust standard errors clustered by the session in parentheses.*** p<0.01,** p<0.05,* p<0.10

Table B7: OLS regression Trust Game by contribution and identity

	Transfer to Leaver		Transfer to Remainer		Transfer to UK	
	(L)	(R)	(L)	(R)	(L)	(R)
Heal	-0.177	-0.108	0.232	0.354	-0.060	0.455*
	(0.307)	(0.304)	(0.362)	(0.328)	(0.314)	(0.234)
Divide	-0.242	-0.397	0.035	0.200	-0.065	0.135
	(0.318)	(0.261)	(0.310)	(0.245)	(0.284)	(0.208)
Controls	YES	YES	YES	YES	YES	YES
Constant	1.722	-0.989	2.908	-2.902**	1.899	-0.158
	(1.442)	(1.506)	(4.260)	(1.446)	(1.294)	(1.004)
Observations	239	266	239	266	239	266

Note: Robust standard errors clustered by the session in parentheses.*** p<0.01,** p<0.05,* p<0.10

Table B8: OLS regression Public Goods Game by contribution and identity

	Transfer to Leaver		Transfer to Remainer		Transfer to UK	
	(L)	(R)	(L)	(R)	(L)	(R)
Heal	0.055	-0.252	0.796**	-0.625**	0.435	-0.102
	(0.338)	(0.312)	(0.369)	(0.276)	(0.294)	(0.253)
Divide	-0.533	-0.343	0.992***	-0.432*	0.199	-0.141
	(0.352)	(0.313)	(0.352)	(0.237)	(0.260)	(0.226)
Controls	YES	YES	YES	YES	YES	YES
Constant	3.009**	0.813	3.868**	1.853	2.990***	0.966
	(1.491)	(1.864)	(1.828)	(1.396)	(1.148)	(1.269)
Observations	239	266	239	266	239	266

Note: Robust standard errors clustered by the session in parentheses.*** p<0.01,** p<0.05,* p<0.10

Table B9: Modified Public Goods Game by contribution and identity

	Transfer to Leaver		Transfer to Remainer		Transfer to UK	
	(L)	(R)	(L)	(R)	(L)	(R)
Heal	0.395	-0.192	0.594**	-0.446	0.558*	0.355
	(0.285)	(0.373)	(0.292)	(0.297)	(0.286)	(0.301)
Divide	-0.005	-0.444	0.303	0.134	0.382	0.342
	(0.286)	(0.360)	(0.307)	(0.248)	(0.290)	(0.272)
Controls	YES	YES	YES	YES	YES	YES
Constant	2.563**	0.069	5.091***	1.153	5.466**	-0.691
	(1.168)	(2.151)	(1.931)	(1.459)	(2.249)	(1.392)
Observations	239	266	239	266	239	266

Note: Robust standard errors clustered by the session in parentheses.*** p<0.01,** p<0.05,* p<0.10

B.3.2 Alternative analysis of social norm behaviours

We find differences between individual and social norms if the individual is a Leaver who was in the *Divide* treatment ($p = 0.004$). However, Remainers evaluations of vandalising a Remainers garden differs if the participant was in the *Control* and *Divide* treatments ($p = 0.050$ and $p = 0.033$, respectively). In both cases, the Remainers personal norm is further skewed towards very socially inappropriate than the in-group norm. Thus we observe discrepancies in norm evaluations of Remainers between personal and group norms, whereas those in the *Heal* treatment display greater norm conformity. Regarding the family marriage scenario, we find differences between a Leavers personal norm and in-group norm in the *Heal* and *Divide* treatments ($p = 0.000$, respectively). We observe a similar pattern for Remainers in the Family scenario across all treatments ($p = 0.000$). The scenario regarding reporting bribery has no statistical difference between personal and group norms, except for Remainers in the *Divide* treatment ($p = 0.004$) - the personal norm is more skewed to very socially appropriate than the (in-)group norm. For punishing bribery, we observe differences between personal and group norms for Leavers in the *Heal* and *Divide* treatments

($p = 0.024$ and $p = 0.096$, respectively). For Remainers evaluation of punishing bribery we observe divergence in the *Control* and *Divide* treatments ($p = 0.007$ and $p = 0.004$).

Table B10 presents the disaggregated norm effects on cooperative behaviour. The dependent variable is the pooled mean contributions for all receivers (Leavers, Remainers and UK) for each game, respectively. We add demographic controls and the elicited beliefs. The control variables include age, gender, risk preferences, religion, employment status, ethnicity, educational attainment and marital status. The (Own) refers to individuals' norm evaluation. The (Leaver) refers to the majority of Leaver's perceptions. The (Remainer) refers to the majority of Remainer's perceptions. We observe a limited effect of disaggregated norm-compliance. More consistently, we observe that individuals norm perceptions impacts behaviour.

Table B10: OLS Regression on effects of social norms on mean contributions in games

Dep. variable: Mean amount sent	PD	DG	TG	mPGG	PGG
	(1)	(2)	(3)	(4)	(5)
Leaver(=1, Remainer=0)	-0.038	-0.097	-0.182	0.138	0.066
	(0.205)	(0.191)	(0.214)	(0.187)	(0.193)
normvandl(Own)	0.045	0.582*	0.545	-0.051	0.485
	(0.366)	(0.340)	(0.414)	(0.299)	(0.346)
normvandl(Leaver)	0.448	-0.175	-0.131	-0.217	-0.287
	(0.348)	(0.267)	(0.220)	(0.185)	(0.245)
normvandl(Remainer)	0.058	0.043	0.158	0.055	-0.114
	(0.177)	(0.160)	(0.155)	(0.151)	(0.196)
normvandr(Own)	-0.147	-0.065	0.224	0.191	0.073
	(0.336)	(0.380)	(0.453)	(0.234)	(0.301)
normvandr(Leaver)	-0.048	-0.181	-0.165	-0.088	0.249
	(0.181)	(0.166)	(0.160)	(0.169)	(0.193)
normvandr(Remainer)	0.150	-0.331	-0.354	-0.261	-0.062
	(0.261)	(0.206)	(0.234)	(0.204)	(0.223)
normfamily(Own)	0.496**	0.461***	0.170	0.217	0.306*
	(0.221)	(0.170)	(0.155)	(0.169)	(0.182)
normfamily(Leaver)	-0.554**	-0.103	0.101	-0.252	-0.149

Continued on next page

Table B10 – continued from previous page

Dep. variable: Mean amount sent	PD	DG	TG	mPGG	PGG
	(1)	(2)	(3)	(4)	(5)
	(0.221)	(0.166)	(0.154)	(0.161)	(0.217)
normfamily(Remainer)	-0.110	-0.129	-0.179	-0.117	-0.135
	(0.209)	(0.179)	(0.154)	(0.167)	(0.215)
normbribe1(Own)	-0.156	-0.150	0.124	-0.076	0.098
	(0.338)	(0.270)	(0.303)	(0.310)	(0.345)
normbribe1(Leaver)	0.088	-0.036	0.102	0.004	0.120
	(0.254)	(0.216)	(0.175)	(0.229)	(0.236)
normbribe1(Remainer)	-0.056	0.134	-0.148	0.345	-0.138
	(0.273)	(0.262)	(0.276)	(0.271)	(0.308)
normbribe2(Own)	0.268	0.186	0.142	0.212	-0.285
	(0.361)	(0.325)	(0.265)	(0.351)	(0.322)
normbribe2(Leaver)	-0.430*	-0.046	-0.148	-0.282	0.135
	(0.253)	(0.216)	(0.206)	(0.215)	(0.252)
normbribe2(Remainer)	0.120	-0.069	0.196	0.025	0.191
	(0.351)	(0.280)	(0.237)	(0.276)	(0.259)
Controls	YES	YES	YES	YES	YES
Constant	5.029***	1.625	0.658	1.090	0.914
	(1.152)	(1.032)	(1.051)	(0.938)	(0.989)
Observations	504	505	505	505	505

Note: Robust standard errors clustered by the session in parentheses.*** p<0.01,** p<0.05,* p<0.10

C

Appendix C for Chapter 4

Appendix C.1 - Instructions

C.1.1 Instructions for Standard Mind Game treatment (UK)

Please think of a number from 1 to 10 (1, 2, 3, 4, 5, 6, 7, 8, 9, and 10). Please keep this number in your mind.

A random number will be shown on the screen. The number will either be 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 (any number could appear).

If the number on the screen is the same as the one you have in mind, you will get one point. If the number on the screen is not the number you have in your mind then you will not get a point. For example, if you choose '3', and the screen shows '3', you get 1 point. If not, then you will earn zero points.

You will repeat this game 10 times.

Please indicate if the number you chose matches the number shown (Please think of your number before pressing 'next'). Please Click 'YES' if the numbers are the same or 'NO' if the numbers are different [1].

C.1.2: Instructions for modified Mind Game treatments

You can earn up to 10 additional points for a [Ndebele or Shona and Leaver or Remainer] participant. **You will receive 4 points for doing this task**.

Please think of a number from 1 and 10 (1, 2, 3, 4, 5, 6, 7, 8, 9, and 10) in this scenario. Please keep this number in your mind.

A random number will be shown on the screen. The number will either be 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 (any number could appear).

If the number on the screen is the same as the one you have in mind, a [Ndebele/Shona and Leaver/Remainer] player will get a point. If the number on the screen is not the number you have in your mind then they [Ndebele/Shona and Leaver/Remainer] will not get a point. For example, if you chose '3', and the screen shows '3', a Ndebele player will get 1 point. If not, then they will earn zero.

You will repeat this game 10 times.

Please indicate if the number you chose matches the number shown (Please think of your number before pressing 'next'). Please Click 'YES' if the numbers are the same or 'NO' if the numbers are different.

[1] Instructions for Zimbabwean experiments are unchanged.

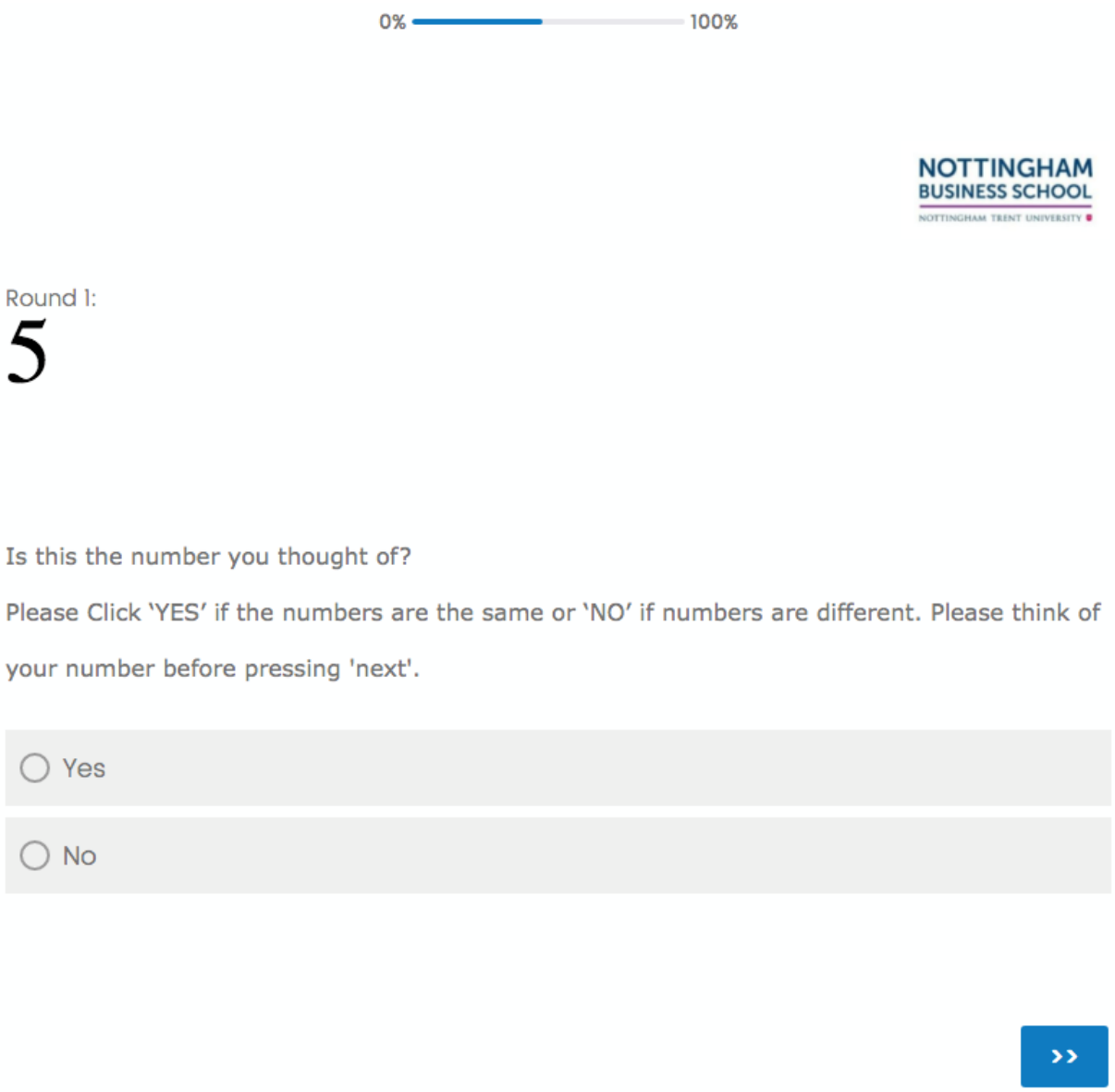


Figure C.1: Example question

C.1.3 Instructions for social norms (UK)

New Task

In this task, you will receive a description of a scenario.

You will be asked to consider each action. Please decide, for each of the possible actions, whether taking that action would be "socially appropriate" or "socially inappropriate".

Please consider your thoughts in each scenario and what you think the actions of Leavers and Remainers are.

In other words, you will be asked to consider each scenario three times:

By socially appropriate, we mean actions that [you or Leavers or Remainers] think is the "correct" thing to do. By socially inappropriate, we mean actions that [you or Leavers or Remainers] would agree is "not the correct" thing to do.

We would like you to answer as truthfully as possible in each of your responses.

After this, your response is compared with the answer that most Leave, and Remain participants chose. If your evaluation is the same as the most common answer, you will receive 5 points for this task; otherwise, you will earn 0 points.

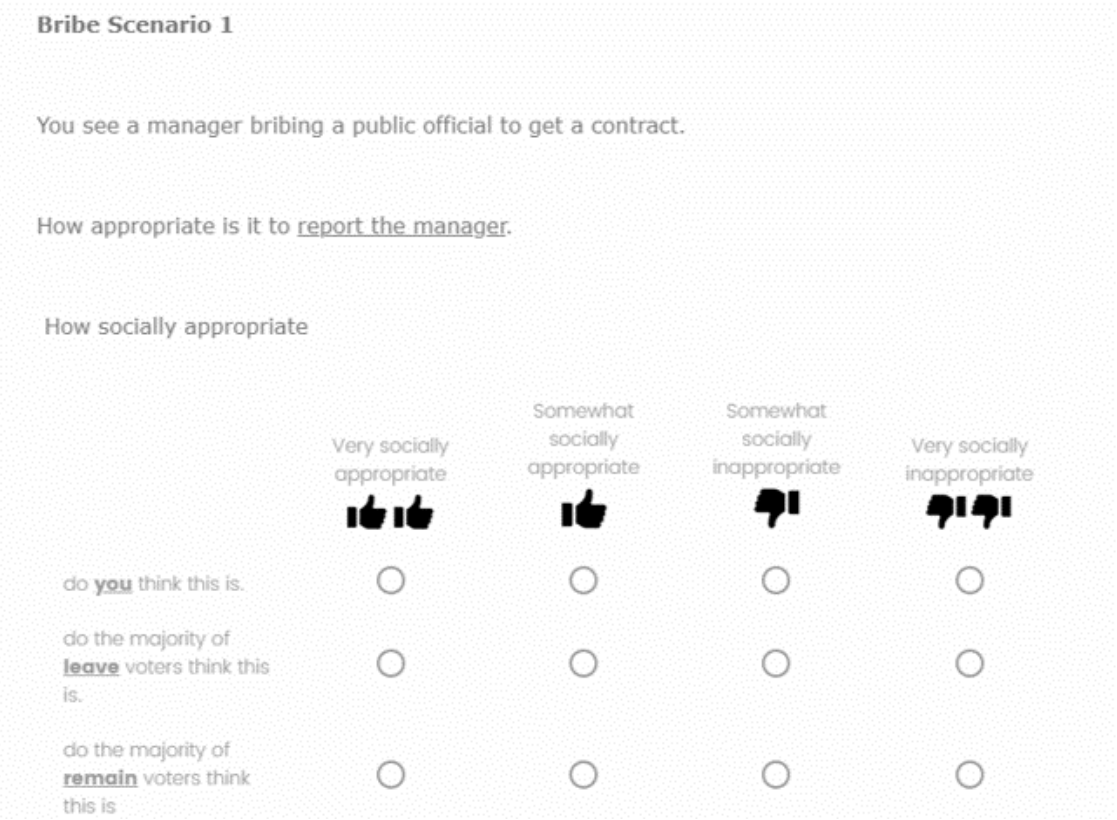


Figure C1: Reporting bribery scenario (UK)

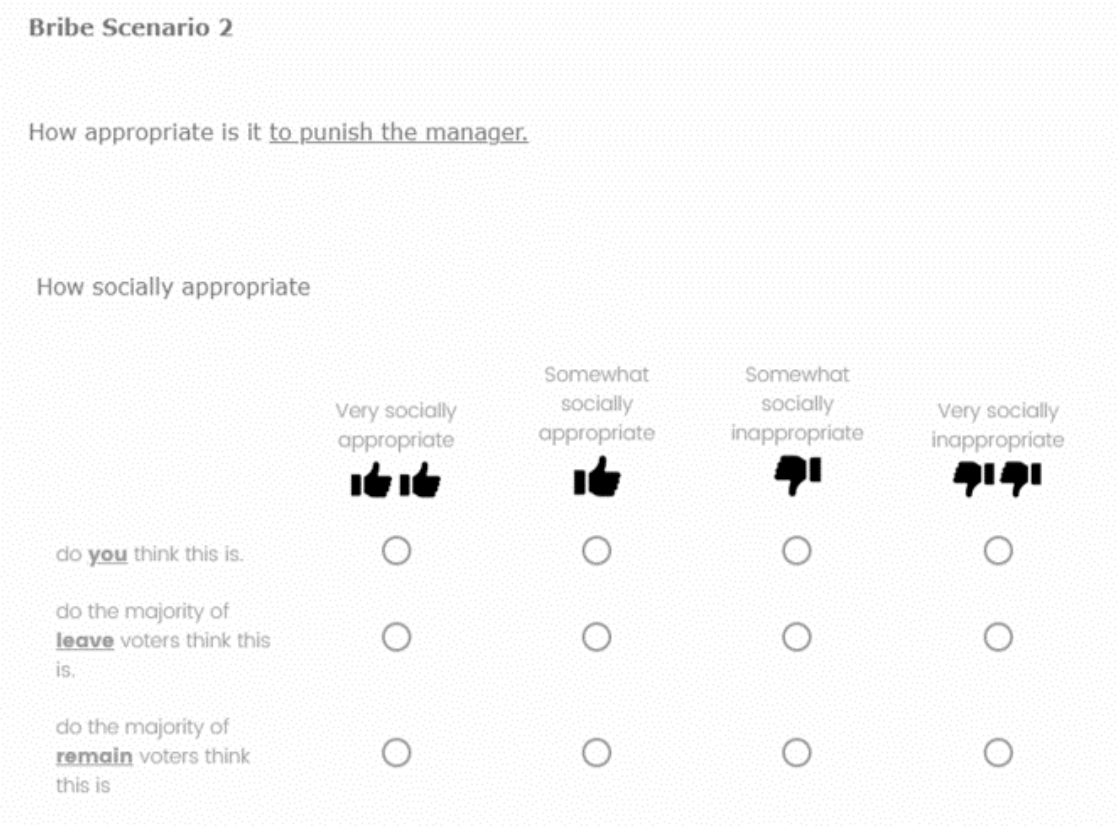


Figure C2: Punishing bribery scenario (UK)

C.1.4 Instructions for social norms (Zimbabwe)

New Task

In this task, you will receive a description of a scenario.

You will be asked to consider each action. For each of the possible actions, you must decide whether taking that action would be "socially appropriate" or "socially inappropriate".

You will be given three options:

By socially appropriate, we mean behaviour that you think is the "correct" thing to do. By socially appropriate, we mean behaviour that most people in your ethnicity would agree is the "correct" thing to do By socially appropriate, we mean behaviour that most people of the other ethnicity would agree is the "correct" thing to do.

In each of your responses, we would like you to answer as truthfully as possible, based on your opinions of what is socially appropriate or socially inappropriate behaviour.

After this, your response is compared with the average response of what most Shona and Ndebele participants chose. If your evaluation is the same as this average, you will receive 5 points for this task; otherwise, you will receive zero.

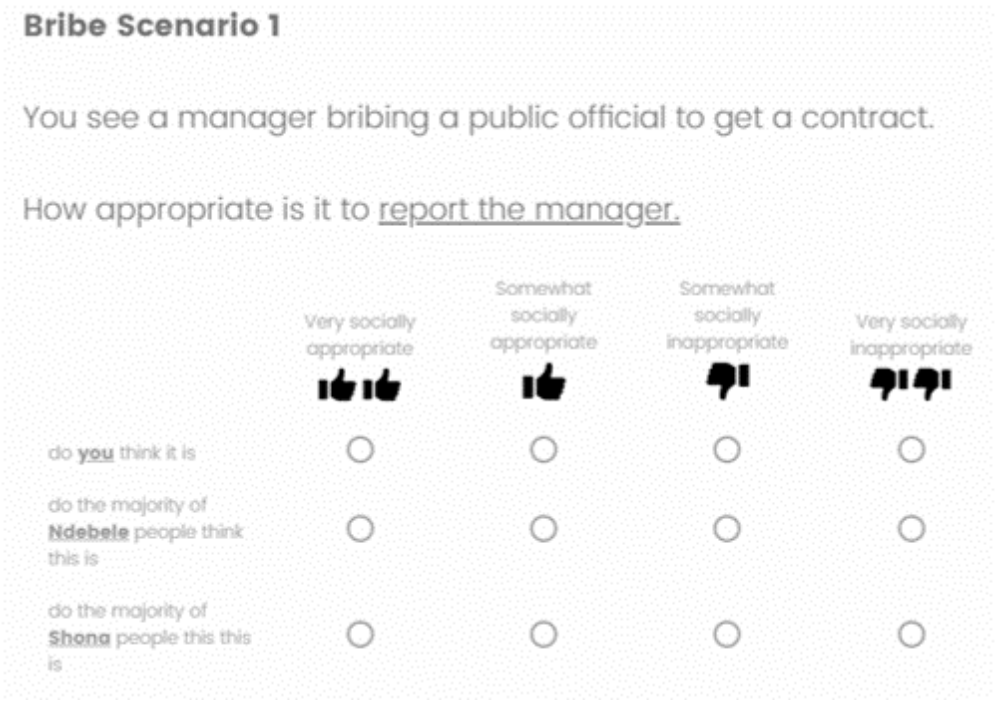


Figure C3: Reporting bribery scenario (Zimbabwe)

Figure C4: Punishing bribery scenario (Zimbabwe)

Appendix C.2 - Additional data analysis

C.2.1 Random effects

To account for participant's repeated decision-making over ten rounds, we also run panel random fixed-effects in both the UK and Zimbabwe, see Table C1 and Table C2. The dependent variable is the "Yes" reports in each treatment. We also report the marginal effects. We include demographic controls; age, gender, risk preferences, religion, employment status, ethnicity, educational attainment and marital status. We observe qualitatively similar results using random fixed effects as found in the main analysis of this paper.

Table C1: Random fixed-effects regression of dishonesty (UK)

	Standard		Modified (Leaver)		Modified (Remainer)		Modified (British)	
	(1)	(1m)	(2)	(2m)	(3)	(3m)	(4)	(4m)
Leaver (=1, Remainer=0)	0.778**	0.165**	0.079	0.017	0.260	0.061	0.044	0.009
	(0.364)	(0.079)	(0.264)	(0.057)	(0.225)	(0.051)	(0.202)	(0.040)
govtrust	0.004	0.001	-0.008	-0.002	0.017	0.004	-0.007	-0.001
	(0.052)	(0.010)	(0.041)	(0.009)	(0.045)	(0.010)	(0.029)	(0.006)
govtrust19	-0.092	-0.019	-0.027	-0.006	0.092*	0.021	0.141***	0.028***
	(0.069)	(0.014)	(0.046)	(0.010)	(0.056)	(0.013)	(0.043)	(0.008)
Polarisation								
proud	-0.200**	-0.040**	-0.152**	-0.032**	-0.095	-0.022	-0.073*	-0.014*
	(0.082)	(0.017)	(0.075)	(0.016)	(0.070)	(0.017)	(0.042)	(0.008)
winner	-1.166***	-0.236***	0.220	0.047	0.085	0.020	-0.165	-0.033
	(0.378)	(0.076)	(0.266)	(0.056)	(0.299)	(0.070)	(0.140)	(0.027)
diffvote	-0.648***	-0.131***	0.003	0.001	-0.376	-0.087	-0.030	-0.006
	(0.226)	(0.043)	(0.207)	(0.044)	(0.240)	(0.060)	(0.125)	(0.025)
align	0.557	0.113	-0.025	-0.005	-0.567**	-0.132**	-0.285	-0.056
	(0.350)	(0.071)	(0.195)	(0.041)	(0.245)	(0.056)	(0.178)	(0.034)
risk	-0.039	-0.008	0.001	0.000	-0.053	-0.012	-0.043**	-0.009**
	(0.043)	(0.009)	(0.056)	(0.012)	(0.037)	(0.009)	(0.020)	(0.004)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Constant	3.280***		0.426		-1.030		-1.813***	
	(1.184)		(0.713)		(0.873)		(0.521)	

Continued on next page

Table C1 – continued from previous page

	Standard		Modified (Leaver)		Modified (Remainer)		Modified (British)	
	(1)	(1m)	(2)	(2m)	(3)	(3m)	(4)	(4m)
Observations	330	330	470	470	520	520	410	410

Notes: Robust standard errors in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table C2: Random fixed-effects regression of dishonesty with margins (Zimbabwe)

	Standard		Modified (Nd)		Modified (Sh)		Modified (Zim)	
	(1)	(1m)	(2)	(2m)	(3)	(3m)	(4)	(4m)
Shona(=1, Ndebele=0)	0.280	0.101	-0.109	-0.040	0.454	0.146	0.086	0.026
	(0.282)	(0.100)	(0.362)	(0.134)	(0.436)	(0.129)	(0.317)	(0.094)
govtrust	-0.016	-0.006	0.023	0.009	0.055	0.019	-0.002	-0.001
	(0.044)	(0.016)	(0.062)	(0.023)	(0.049)	(0.017)	(0.050)	(0.015)
govrisk	0.078	0.029	-0.013	-0.005	-0.018	-0.006	0.060	0.018
	(0.062)	(0.023)	(0.049)	(0.018)	(0.049)	(0.017)	(0.045)	(0.014)
Self-identify (BAse = African)								
Ndebele	0.615	0.240	-0.447	-0.141	0.043	0.017	-0.612	-0.111*
	(0.408)	(0.158)	(0.432)	(0.117)	(0.378)	(0.146)	(0.441)	(0.060)
Shona	0.928***	0.357***	0.120	0.044	-0.528*	-0.179**	1.444***	0.518***
	(0.297)	(0.105)	(0.332)	(0.124)	(0.272)	(0.082)	(0.319)	(0.108)
Zimbabwean	-0.087	-0.031	0.055	0.020	-0.497	-0.170*	0.443**	0.135**
	(0.254)	(0.090)	(0.211)	(0.077)	(0.316)	(0.101)	(0.181)	(0.055)
Shona align	-0.022	-0.008	0.043	0.016	0.112	0.039	-0.134	-0.041
	(0.134)	(0.050)	(0.132)	(0.048)	(0.216)	(0.074)	(0.097)	(0.030)
Ndebele align	0.139	0.051	0.107	0.039	0.267	0.092	0.229	0.071
	(0.111)	(0.041)	(0.158)	(0.057)	(0.171)	(0.059)	(0.156)	(0.048)
risk	-0.029	-0.011	0.055	0.020	0.164***	0.057***	-0.104**	-0.032**
	(0.057)	(0.021)	(0.050)	(0.018)	(0.058)	(0.020)	(0.053)	(0.016)
Controls	YES	YES	YES	YES	YES	YES	YES	YES

Continued on next page

Table C2 – continued from previous page

	Standard		Modified (Nd)		Modified (Sh)		Modified (Zim)	
	(1)	(1m)	(2)	(2m)	(3)	(3m)	(4)	(4m)
Constant	-0.205		-2.511***		-4.102***		-1.508**	
	(0.973)		(0.910)		(1.423)		(0.740)	
Observations	500	500	510	510	490	490	470	470

Notes: Robust standard errors in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

C.2.2 Count model

As our response variable is a count variable, each subject has the same length of observations. We also do not find over dispersion; therefore use the Poisson regression model over negative binomial or zero-inflated. The dependent variable is the fraction of "Yes" reports in each treatment, respectively. We include demographic controls, the control variables include age, gender, risk preferences, religion, employment status, ethnicity, educational attainment and marital status. We anticipate that the more socially appropriate reporting and punishing bribery is perceived, the reduced likelihood of reporting "Yes". Table C3 reports Poisson regression for the experiments conducted in the UK. We also report the marginal effects, which provide the effects in count scales. The results are unchanged. Table C4 reports the Poisson regression for the experiments conducted in Zimbabwe. We observe differences in the count of dishonesty over the 10 rounds in the experiments conducted in Zimbabwe, with more norm effects over the treatments. However, the self-identification results remain unchanged.

Table C3: Poisson regression of dishonesty (UK)

	Standard		Modified (Leaver)		Modified (Remainer)		Modified (British)	
	(1)	(1m)	(2)	(2m)	(3)	(3m)	(4)	(4m)
0.identity_new	1.028	0.055	0.188	0.018	0.592	0.081	-0.133	-0.016
	(0.870)	(0.049)	(0.287)	(0.028)	(0.428)	(0.060)	(0.288)	(0.035)
Norms (1 = majority socially appropriate)								
bribe1 (Leaver)	-1.198	-0.059	-2.004***	-0.188***	0.608	0.081	0.016	0.002
	(0.755)	(0.036)	(0.502)	(0.034)	(1.008)	(0.134)	(0.634)	(0.078)
bribe1 (Remainer)	-16.574***	-0.810***	-0.846	-0.079	1.736***	0.232***	-0.097	-0.012
	(2.188)	(0.128)	(0.844)	(0.067)	(0.557)	(0.081)	(0.580)	(0.071)
bribe1 (Own)	18.026***	0.881***	3.488***	0.327***	-2.166*	-0.289*	0.211	0.026
	(2.584)	(0.140)	(1.135)	(0.058)	(1.190)	(0.160)	(0.366)	(0.046)
bribe2 (Leaver)	0.654	0.032	-2.752***	-0.258***	-2.421***	-0.323***	0.490	0.060
	(0.783)	(0.037)	(0.631)	(0.056)	(0.383)	(0.052)	(0.626)	(0.075)
bribe2 (Remainer)	0.277	0.014	-12.546***	-1.177***	2.011***	0.269***	-1.314***	-0.161***
	(1.442)	(0.070)	(1.232)	(0.272)	(0.227)	(0.047)	(0.400)	(0.042)
bribe2 (Own)	-2.409	-0.118	14.162***	1.329***	-0.434	-0.058	0.876	0.107
	(2.422)	(0.110)	(1.260)	(0.286)	(0.644)	(0.086)	(0.690)	(0.084)
vote_proud	-0.053	-0.003	-0.267***	-0.025***	0.027	0.004	-0.151**	-0.019**
	(0.110)	(0.005)	(0.081)	(0.007)	(0.087)	(0.012)	(0.065)	(0.008)

Continued on next page

Table C3 – continued from previous page

	Standard		Modified (Leaver)		Modified (Remainer)		Modified (British)	
	(1)	(1m)	(2)	(2m)	(3)	(3m)	(4)	(4m)
winner	-0.390	-0.019	-0.038	-0.004	0.256	0.034	-0.375	-0.046
	(0.784)	(0.037)	(0.389)	(0.037)	(0.450)	(0.060)	(0.372)	(0.044)
diff vote	0.084	0.004	0.330	0.031	-0.305	-0.041	-0.322	-0.039
	(0.295)	(0.015)	(0.270)	(0.027)	(0.224)	(0.030)	(0.250)	(0.030)
align	-0.427	-0.021	0.102	0.010	-0.892***	-0.119***	-0.485	-0.059*
	(0.376)	(0.017)	(0.247)	(0.024)	(0.268)	(0.032)	(0.300)	(0.034)
Constant	2.667		-2.407		-0.428		-3.703*	
	(3.841)		(1.675)		(1.598)		(2.190)	
Observations	330	330	470	470	520	520	410	410

Notes: Robust standard errors in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table C4: Poisson regression of dishonesty (Zimbabwe)

	Standard		Modified (Ndebele)		Modified (Shona)		Modified (Zim)	
	(1)	(1m)	(2)	(2m)	(3)	(3m)	(4)	(4m)
2.identity	0.307	0.093	0.063	0.022	0.750	0.185	-0.210	-0.053
	(0.193)	(0.059)	(0.101)	(0.035)	(0.924)	(0.201)	(0.358)	(0.096)
Norms (1 = majority socially appropriate)								
bribe1 (Nd)	0.196	0.062	-0.742***	-0.228***	-0.277	-0.077	-0.147	-0.035
	(0.298)	(0.095)	(0.160)	(0.044)	(0.219)	(0.061)	(0.521)	(0.122)
bribe1 (Sh)	-0.657***	-0.201***	0.359*	0.127*	0.954***	0.303***	0.046	0.011
	(0.219)	(0.059)	(0.193)	(0.065)	(0.246)	(0.095)	(0.162)	(0.039)
bribe1 (Own)	0.347	0.114	0.459	0.168	-0.307	-0.087	-0.444	-0.106
	(0.330)	(0.108)	(0.298)	(0.109)	(0.215)	(0.064)	(0.282)	(0.069)
bribe2 (Nd)	-0.722***	-0.296***	0.166*	0.056*	-0.429	-0.135	0.057	0.013
	(0.164)	(0.082)	(0.094)	(0.033)	(0.265)	(0.095)	(0.579)	(0.133)
bribe2 (Sh)	0.313	0.093	-0.551**	-0.219*	0.131	0.035	-0.279	-0.071
	(0.207)	(0.057)	(0.256)	(0.119)	(0.464)	(0.121)	(0.369)	(0.102)
bribe2 (Own)	0.469***	0.127***	-0.179	-0.067	-0.274	-0.087	-0.350	-0.096
	(0.116)	(0.027)	(0.203)	(0.083)	(0.728)	(0.262)	(0.297)	(0.093)
Self-identity(Base=African)								
Ndebele	0.342*	0.143			-0.154	-0.055	-0.885***	-0.112***
	(0.194)	(0.089)			(0.234)	(0.086)	(0.222)	(0.017)

Continued on next page

Table C4 – continued from previous page

	Standard		Modified (Ndebele)		Modified (Shona)		Modified (Zim)	
	(1)	(1m)	(2)	(2m)	(3)	(3m)	(4)	(4m)
Shona	0.927***	0.538***	-0.331***	-0.089***	-0.290	-0.097	1.235***	0.464**
	(0.200)	(0.133)	(0.082)	(0.024)	(0.597)	(0.171)	(0.339)	(0.224)
Zimbabwean	-0.395*	-0.115*	0.280	0.102	-0.733	-0.200	0.417***	0.098***
	(0.210)	(0.060)	(0.311)	(0.120)	(0.469)	(0.130)	(0.132)	(0.035)
shona align	-0.138*	-0.044*	-0.231**	-0.081**	-0.008	-0.002	-0.052	-0.012
	(0.077)	(0.025)	(0.111)	(0.039)	(0.353)	(0.099)	(0.077)	(0.019)
ndebele align	0.029	0.009	-0.113	-0.040	0.547**	0.154**	0.314	0.075
	(0.114)	(0.036)	(0.142)	(0.050)	(0.266)	(0.073)	(0.388)	(0.091)
Constant	-3.853***		-2.169**		-6.265***		-1.567***	
	(0.467)		(1.045)		(1.870)		(0.568)	
Observations	440	440	450	450	440	440	460	460

Notes: Robust standard errors in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

D

Appendix D for Chapter 5

Appendix D.1 - Instructions

The purpose of the research is to study how individuals make decisions in a number of economically relevant tasks and their beliefs about how others make decisions in those same tasks. You will be presented with several economic scenarios in which you are tasked to make a value judgement and a decision on the social appropriateness of different scenarios. As you do this, you will receive points, depending on the response you have provided to the survey.

You will earn a £2.50 participation fee (simply for participating). In addition, you may also earn extra money depending on your decisions and the choices of others in the study. You will make decisions using points. These points will be converted into money. The task will take approximately 20 mins.

The computer will be used to work out the amounts of payment each of you gets back at the end. At the end of the study, the computer will randomly select one of your decisions for

payout. This will decide your final payment. You will not know until the end of the study which task you will be paid for; please make your decisions in each task carefully. You will not receive feedback on the outcome of any tasks. After the tasks are finished, you will also be asked to complete a questionnaire.

The anonymity of your responses to all parts of the exercise (the tasks and questionnaire) is guaranteed.

Thank you, and please continue.

Big Five questionnaire

In general, I see myself as a person who:

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
Worries a lot	○	○	○	○	○
Gets nervous easily	○	○	○	○	○
Is relaxed, handles stress well	○	○	○	○	○
Is talkative	○	○	○	○	○
Is outgoing, sociable	○	○	○	○	○
Is reserved	○	○	○	○	○
Is original, comes up with new ideas	○	○	○	○	○
Values artistic, aesthetic experiences	○	○	○	○	○
Has an active imagination	○	○	○	○	○
Is sometimes rude to others	○	○	○	○	○
Has a forgiving nature	○	○	○	○	○
Is considerate and kind to almost everyone	○	○	○	○	○
Does a thorough job	○	○	○	○	○
Tends to be lazy	○	○	○	○	○
Does things efficiently	○	○	○	○	○

Figure D1: Big Five Inventory Questions

Appendix D.2 - Supplementary data analysis

D.2.1 Non-linear relationship of personality traits

Following Ben-Ner & Kramer (2011), we also included the squared terms of personality traits. To identify if the relationship between aggregated cooperativeness and personalities are non-linear. We find that there is a non-linear relationship with Extraversion. Participants higher in extraversion give significantly more in the Dictator Game, but at higher levels, the results

suggest that the relationship becomes negative. The finding implies that extraverts may be willing to be cooperative to an extent in purely altruistic games, such as the Dictator Game.

Table D1: Regression for personality effects in behaviour

	PD	DG	TG	PGG	mPGG
	(1)	(2)	(3)	(4)	(5)
Leaver	0.915	-0.332	-0.320	0.733	-0.962
	(1.607)	(1.422)	(1.385)	(1.305)	(1.554)
Neuroticism	0.869	0.877	-0.804	-0.215	-0.647
	(1.034)	(0.997)	(1.260)	(0.911)	(0.988)
Neuroticism x Leaver	-0.298	-0.030	-0.084	-0.170	0.103
	(0.261)	(0.221)	(0.250)	(0.210)	(0.228)
Extraversion	-0.774	2.102*	-0.750	0.257	-0.193
	(0.979)	(1.246)	(1.006)	(0.741)	(1.040)
Extraversion x Leaver	-0.134	0.077	0.000	-0.458**	-0.303
	(0.270)	(0.286)	(0.230)	(0.212)	(0.263)
Openness	-0.989*	-0.525	-0.034	-0.072	0.447
	(0.568)	(0.442)	(0.505)	(0.471)	(0.554)
Openness x Leaver	0.248	0.299**	0.307**	0.042	-0.013
	(0.193)	(0.152)	(0.146)	(0.146)	(0.165)
Agreeableness	-0.404	-0.354	1.267	0.700	0.866
	(0.971)	(0.862)	(1.128)	(0.684)	(0.956)
Agreeableness x Leaver	-0.051	-0.119	-0.225	0.105	0.046
	(0.297)	(0.252)	(0.268)	(0.217)	(0.277)
Conscientiousness	0.553	2.517	-1.900	-0.512	0.854
	(1.399)	(1.553)	(1.707)	(1.618)	(1.810)
Conscientiousness x Leaver	-0.097	-0.121	0.066	0.237	0.430
	(0.302)	(0.292)	(0.307)	(0.291)	(0.314)
$Neuroticism^2$	-0.093	-0.130	0.111	0.027	0.101
	(0.162)	(0.150)	(0.196)	(0.139)	(0.150)
$Extraversion^2$	0.137	-0.348*	0.098	0.003	0.049
	(0.153)	(0.192)	(0.156)	(0.116)	(0.167)
$Openness^2$	0.119	0.053	-0.007	-0.005	-0.067

Continued on next page

Table D1 – continued from previous page

	PD	DG	TG	PGG	mPGG
	(1)	(2)	(3)	(4)	(5)
	(0.076)	(0.062)	(0.071)	(0.068)	(0.078)
$Agreeableness^2$	0.052	0.073	-0.154	-0.148	-0.136
	(0.140)	(0.132)	(0.165)	(0.101)	(0.144)
$Conscientiousness^2$	-0.045	-0.333	0.233	0.084	-0.106
	(0.191)	(0.205)	(0.229)	(0.213)	(0.246)
Controls	YES	YES	YES	YES	YES
Constant	2.829	-5.847	5.376	1.370	-2.110
	(3.315)	(4.004)	(4.606)	(3.655)	(3.852)
N	504	505	505	505	505

Note: The dependent variable is the pooled mean contributions for all receivers (leave, remain, non-voters and UK) in each game. We add demographic controls and beliefs. *** p<0.01, ** p<0.05, * p<0.10.

D.2.2 Each individual 15-item personality trait

As a further robustness check, we test the impact on behaviours by each individual 15-item indicator. The disaggregated 15-item regression shows that Neuroticism reduces cooperation in all social dilemma games, except in the prisoner's dilemma. The negative relationship between Neuroticism and prosocial behaviour has been observed in the psychological and economics literature due to highly neurotic individuals fearing the consequences of cooperation (Sabater-Grande et al. 2022, Evans & Revelle 2008). Low levels of extraversion - being *reserved* increases the likelihood of defection in the prisoner's dilemma but increases cooperation in the public goods game and modified public goods game. Swope et al. (2008), also find a significant positive relationship between extraversion and prosocial behaviour. Surprisingly, Openness through being *original* and having an *active imagination* reduces cooperation in the trust game and public goods game, respectively. However, *artistic* participants are more cooperative in the public goods game. In addition, individuals low in agreeableness who are *sometimes rude* are more likely to defect in the prisoner's dilemma. Surprisingly, individuals with a forgiving nature or agreeable contribute less to the common pot in the standard public goods game but more in the modified public goods game. The

results in the modified public goods game is supported by findings by Kagel & McGee (2014). The negative relationship between agreeableness and prosocial behaviour may be explained by findings in Table D1, which suggests that agreeable individuals are less inclined to be prosocial at higher levels. Individuals low in Conscientiousness who are *lazy* are more likely to coordinate and cooperate in the modified public goods game, potentially due to *lazy* individuals not truly understanding the strategy set (Burger 2015). Whereas more Conscientious individuals coordinate less.

Table D2: Regression for personality effects in behaviour (15-item)

	PD	DG	TG	PGG	mPGG
	(1)	(2)	(3)	(4)	(5)
Leaver	0.031	0.009	0.044	-0.008	-0.068
	(0.197)	(0.152)	(0.176)	(0.156)	(0.167)
Neuroticism					
worries a lot	0.209**	0.062	-0.160*	-0.127*	-0.029
	(0.086)	(0.078)	(0.089)	(0.074)	(0.090)
gets nervous	-0.079	-0.138**	-0.029	-0.026	-0.056
	(0.098)	(0.069)	(0.081)	(0.072)	(0.086)
is relaxed	0.048	0.039	-0.086	-0.110	-0.056
	(0.090)	(0.086)	(0.092)	(0.072)	(0.093)
Extraversion					
is talkative	0.014	-0.034	-0.076	0.071	0.001
	(0.087)	(0.077)	(0.072)	(0.067)	(0.073)
is outgoing	-0.030	-0.066	-0.000	-0.008	0.021
	(0.099)	(0.080)	(0.082)	(0.075)	(0.080)
is reserved	-0.201**	0.029	0.101	0.126*	0.162*
	(0.087)	(0.075)	(0.072)	(0.073)	(0.083)
Openness					
is original	-0.060	-0.061	-0.156**	-0.007	-0.051
	(0.087)	(0.072)	(0.074)	(0.073)	(0.081)
values artistic	0.084	0.120	0.118	0.150**	0.104
	(0.080)	(0.085)	(0.083)	(0.069)	(0.084)
active imagination	-0.085	-0.041	0.134	-0.141**	0.002

Continued on next page

Table D2 – continued from previous page

	PD	DG	TG	PGG	mPGG
	(1)	(2)	(3)	(4)	(5)
	(0.086)	(0.075)	(0.082)	(0.071)	(0.088)
Agreeableness					
sometimes rude	-0.140**	-0.040	0.026	-0.006	-0.054
	(0.066)	(0.062)	(0.064)	(0.057)	(0.069)
forgiving nature	0.069	0.071	0.055	-0.100*	-0.006
	(0.080)	(0.067)	(0.066)	(0.060)	(0.076)
is considerate	-0.069	0.015	0.097	0.049	0.162*
	(0.112)	(0.092)	(0.096)	(0.083)	(0.093)
Conscientiousness					
thorough job	0.125	-0.089	-0.093	-0.051	0.005
	(0.113)	(0.091)	(0.100)	(0.094)	(0.110)
lazy	0.146**	0.093	-0.029	0.085	0.144*
	(0.071)	(0.066)	(0.056)	(0.064)	(0.074)
efficient	-0.027	-0.168**	-0.101	-0.088	-0.029
	(0.113)	(0.084)	(0.095)	(0.093)	(0.102)
vote proud	-0.014	-0.034	-0.125***	0.012	-0.018
	(0.051)	(0.048)	(0.047)	(0.041)	(0.050)
winner	-0.274	-0.294*	-0.185	-0.275*	-0.138
	(0.195)	(0.150)	(0.144)	(0.151)	(0.153)
align	0.186	-0.037	0.050	0.171	-0.004
	(0.171)	(0.138)	(0.148)	(0.139)	(0.153)
Controls	YES	YES	YES	YES	YES
Constant	1.966	2.384*	1.589	1.749	-1.333
	(1.406)	(1.272)	(1.246)	(1.166)	(1.302)
N	504	505	505	505	505

Note: The dependent variable is the pooled mean contributions for all receivers (leave, remain, non-voters and UK) in each game. We add demographic controls and beliefs. *** p<0.01, ** p<0.05, * p<0.10.

Appendix D.3 - Big 5 personalities and political orientation

D.3.1 Political orientation and cooperativeness

The political orientation measure is constructed by ranking the voting in the 2019 General election. Greater values correspond to greater conservative political orientation. We test aggregated cooperativeness on behaviours given political orientation. The dependent variable is the pooled amounts sent to each receiver (Leaver, Remainer and British) for each social dilemma game. We include demographic controls; age, gender, risk preferences, religion, employment status, ethnicity, educational attainment and marital status. We fail to find compelling evidence of the relationship between personality traits and political orientation. Our results are contrary to findings by Hirsh et al. (2010). However, they are in line with Alford & Hibbing (2007), who also fail to find predictive power of personality traits with political ideology.

Table D1: Pairwise correlation of Big Five Personality traits and voting in the 2019 General election

	Neuro	**Extra**	**Open**	**Agree**	**Consc**
Brexit Party	0.001	0.020	0.029	0.014	0.022
Change UK	0.123	0.106	0.038	0.127	0.123
Conservative Party	-0.049	0.010	-0.036	-0.044	-0.017
Did not vote	-0.023	-0.020	-0.055	0.038	0.083
Green Party	0.019	0.025	0.027	0.002	0.067
Labour Party	0.013	0.036	0.012	0.019	-0.064
Liberal Democrats	-0.023	-0.032	0.038	-0.039	-0.024
Plaid Cymru	0.005	-0.043	-0.031	-0.020	-0.057
Scottish National Party	0.090	-0.064	0.053	0.002	-0.022

Notes: Baseline summary statistics of participant characteristics in the experiment.

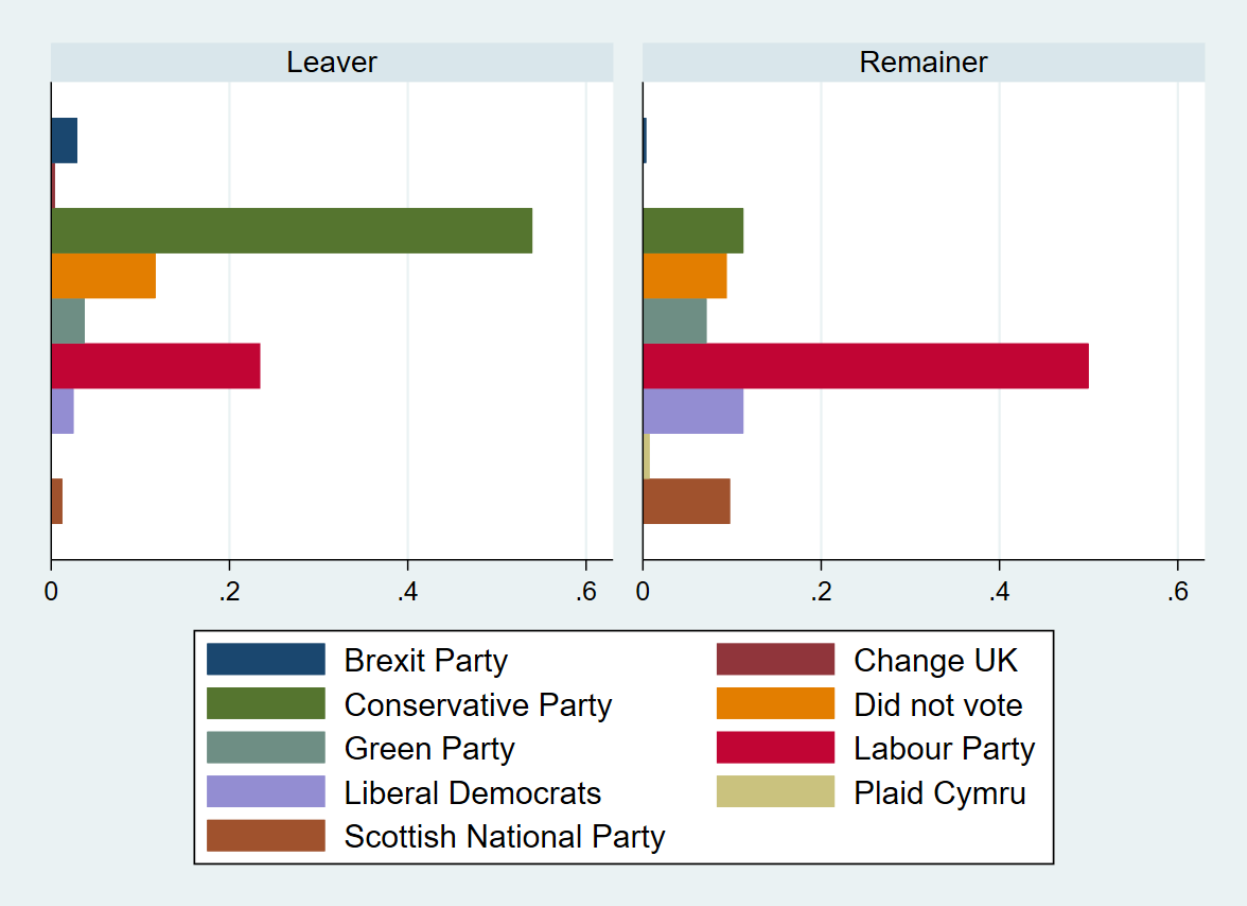


Figure D2: Bar chart of the percentage of Leavers and Remainers who voted in the 2019 General Election

Table D2: Regressions of political orientation and behaviours

	PD	DG	TG	PGG	mPGG
	(1)	(2)	(3)	(4)	(5)
Leaver (=1, Remainer=0)	0.042	-0.028	-0.094	0.152	0.034
	(0.202)	(0.172)	(0.194)	(0.152)	(0.165)
Neuroticism	0.337	-0.219	-0.167	-0.190	-0.029
	(0.249)	(0.213)	(0.239)	(0.212)	(0.229)
pol_orien	-0.385	-0.022	0.085	0.062	0.125
	(0.360)	(0.295)	(0.271)	(0.305)	(0.361)
Neuroticism x pol_orien	-0.062	0.051	-0.024	0.025	0.005
	(0.061)	(0.043)	(0.048)	(0.042)	(0.044)

Continued on next page

Table D2 – continued from previous page

	PD	DG	TG	PGG	mPGG
	(1)	(2)	(3)	(4)	(5)
Extraversion	0.176	-0.031	-0.399*	0.082	-0.023
	(0.255)	(0.289)	(0.232)	(0.214)	(0.299)
Extraversion x pol_orien	-0.048	-0.021	0.055	-0.012	0.022
	(0.057)	(0.052)	(0.048)	(0.040)	(0.052)
Openness	-0.296	-0.136	0.105	0.037	-0.031
	(0.191)	(0.183)	(0.160)	(0.166)	(0.193)
Openness x pol_orien	0.069*	0.031	-0.003	-0.013	0.014
	(0.037)	(0.034)	(0.031)	(0.029)	(0.034)
Agreeableness	-0.344	0.238	0.210	-0.289	0.200
	(0.298)	(0.272)	(0.269)	(0.267)	(0.359)
Agreeableness x pol_orien	0.068	-0.038	0.001	0.038	-0.051
	(0.059)	(0.048)	(0.053)	(0.043)	(0.059)
Conscientiousness	-0.170	-0.037	0.033	0.207	0.332
	(0.312)	(0.302)	(0.287)	(0.323)	(0.440)
Conscientiousness x pol_orien	0.077	-0.005	-0.051	-0.051	-0.024
	(0.065)	(0.058)	(0.061)	(0.056)	(0.072)
Controls	YES	YES	YES	YES	YES
Constant	4.550**	1.548	1.803	1.527	-2.454
	(2.124)	(1.675)	(1.727)	(1.902)	(2.267)
N	452	452	452	452	452

Note: The dependent variable is the pooled mean contributions for all receivers (leave, remain, and the UK) in each game. Political orientation greater values correspond to greater conservative political orientation. We add demographic controls and beliefs. *** p<0.01, ** p<0.05, * p<0.10.

D.3.2 Political orientation and misrepresenting Brexit vote

As a further test, we replicate the regression in the main text, considering misreporting of the Brexit vote and the relationship between political orientation and personality traits. The dependent variable is a dummy variable that equals 1, if participants report the same identity

to the experimenter and online platform, 0 otherwise. In column (2), we include demographic controls. We find evidence that more conservative participants high in Extraversion and Conscientiousness influence the likelihood of misrepresenting voting decisions in the EU referendum, although this is not robust to the inclusion of controls.

Table D3: Probit regression misreporting voting decision and political orientation

	(1)	(2)
Leaver	-0.053	-0.089
	(0.134)	(0.178)
Neuroticism	-0.225	-0.278
	(0.198)	(0.214)
pol_orien	-0.178	-0.103
	(0.292)	(0.318)
Neuroticism x pol_orien	0.038	0.036
	(0.042)	(0.045)
Extraversion	-0.142	-0.070
	(0.219)	(0.230)
Extraversion x pol_orien	0.087*	0.076
	(0.048)	(0.051)
Openness	-0.051	0.005
	(0.151)	(0.162)
Openness x pol_orien	0.011	0.002
	(0.031)	(0.033)
Agreeableness	-0.059	-0.099
	(0.245)	(0.264)
Agreeableness x pol_orien	0.010	0.012
	(0.049)	(0.052)
Conscientiousness	0.421	0.449
	(0.257)	(0.284)
Conscientiousness x pol_orien	-0.094*	-0.091
	(0.055)	(0.059)
Controls	NO	YES

Continued on next page

Table D3 – continued from previous page

	(1)	(2)
Constant	0.154	-3.616**
	(1.468)	(1.840)
N	452	452

Note: The dependent variable is a dummy variable that equals 1, if participants report the same identity to the experimenter and online platform, 0 otherwise. Political orientation greater values correspond to greater conservative political orientation. We add demographic controls and beliefs. *** p<0.01, ** p<0.05, * p<0.10.