Social learning and cooperative behaviour: Evaluating the influence of social learning in social dilemmas

ROBIN WATSON

Social learning and cooperative behaviour: Evaluating the influence of social learning in social dilemmas

Robin Watson (BSc, MA)

A thesis presented for the degree of Doctor of Philosophy

Department of Anthropology
Durham University
February 2021

Table of contents

List of tables

List of figures

Declaration

I confirm that no part of the material presented in this thesis has previously been submitted for a degree in this or any other institution. If material has been generated through joint work, this has been indicated where appropriate. All other sources have been referenced, and quotations suitably indicated.

Statement of copyright

Acknowledgements

Firstly, I'd like to thank Gilbert Roberts and all those at Newcastle University who are the reason I am here writing this in the first place. I'm incredibly grateful for the help I received in putting together the application and the moral support I was given for applying. By extension, thank you to the NINE DTP for funding my project (and for the odd pep talk here and there). Without the support of these people, this project wouldn't have been possible.

To Rachel, Jeremy and Julie. Honestly and sincerely thank you for everything you have done for me throughout this 4-year process. You have been the best supervisors I could possibly have asked for. There isn't any doubt in my mind that I am where I am today because of your support and guidance. You struck the perfect balance between pushing me beyond my comfort zone and providing the reassurance I sometimes needed. I have learned so much during my time as your student and I'm really going to miss working with you all. I'm sure that whatever I go on to do in the future will be made all the better having had the privilege of being your student.

I also need to extend my thanks to Tom at ASU. It was just a stroke of good fortune really that you came over to Durham and I got to meet you, even more so that you gave me the opportunity to travel to Arizona to learn Dallinger. I have no idea what the thesis would have looked like without Dallinger and your continued help and support. You've been really fantastic to work with and the thesis has been made all the better for it. Also, thanks to Matthew. My trip wouldn't have been half as enjoyable (and straightforward) had it not been for you. It was a really great experience being over there and I still rate the Snowbowl as one of the best days out snowboarding, I've ever had.

Thank you to Sally and Sarah who, during my time helping on the DAR module, often provided a friendly ear and some extremely helpful career advice. Thank you to Ninette and the others at Durham Greenspace for your help planning the energy study (which, sadly, COVID cancelled) and recruitment for the field study. Thank you to the behavioural insights team at PHE for giving me the opportunity to complete an extremely interesting internship and particularly Natalie for your ongoing career advice.

Thank you so much to my family. Abbie, you have been supportive right the way through and always been there when I needed you. It’s hard to really describe quite how much of a difference that has made when things have been getting on top of me. I’m really looking forward to what our next steps will be, but I’m sure it will be all the better doing it together. Thank you to Nugget for calming down and not barking all the time. I don’t always need your toys pressed on me but sometimes it’s a welcome break. Thank you to Dad. Whether it’s your help with various bits in the house (interfacing with the roofer comes to mind) or just having a chat through things, your input is always extremely valuable. I don’t think I could have done it without the two of you. It’s sad that mum is no longer around to see what’s going on, I do often wonder what she would make of it all. I suppose she just wouldn’t want a fuss about it. Finally thank you to my friends. Having you all around on Discord has been really great, even if we don’t really chat about what I do. I’m still not so sure about our choice of games to play, but I’ll let that slide. Thank you for being around.

Thesis Abstract

Humans are characterised by a strong proclivity towards two traits: cooperation and social learning. Our ability to cooperate in large unrelated groups and selectively copy and improve upon cultural traits through cumulative cultural evolution are fundamental drivers for our widespread success as a species. However, we know comparatively little regarding the extent that cooperative behaviour spreads through social learning. This thesis focused on three widely studied social learning strategies (payoff bias, conformity and prestige bias) and investigated their influence on cooperative behaviour across three studies. I used online experiments (studies 1 and 2) and an agent-based model (study 1) to assess cooperation (and spite) in economic games and a questionnaire and semi-structured interviews to investigate social learning and cooperation within a real-life context, environmental sustainable behaviour (study 3).

In study 1, model comparison indicated that the changes in cooperation observed across rounds in a cooperative game were best explained by payoff biased social learning than by prestige bias or conformity, though a follow up agent-based model showed that the effect of social learning was small compared to variation within participant's inclination to cooperation. There was additionally some evidence of higher cooperation under a snowdrift payoff structure than under a prisoner's dilemma. Study 2 found marked preferences towards altruism over spiteful behaviour: participants were not influenced by the source (success weighted or frequency weighted) of social information but were slightly more altruistic after viewing altruistic social information. Study 3 found indirect support for payoff bias and conformity within real-life sustainable behaviour decisions, but little support for prestige bias. Individuals reported perceiving strong environmental injunctive norms in their friends and family (conformity) and that the biggest barrier to their sustainable behaviour was the costs of such behaviour (payoff). Sources of social information such as familiar or knowledgeable individuals were generally preferred to conformity or high status weighted social information. Like study 1, this effect was small compared to the influence of individual's personal norms towards sustainability. In addition, experimentally manipulated vignettes framed with payoff, conformity or prestige

information did not encourage individuals to choose a more sustainable over less sustainable participant reward.

Overall, the findings in this thesis indicate a clear hierarchy in social learning strategy use: payoff bias was the most influential, then conformity and then prestige. This generally supported previous research identifying the influence of payoff bias but contrasted with strong theoretical predictions concerning prestige. However, the impact of social learning throughout this thesis was small compared to other influential factors. The thesis is concluded by discussing some methodological and theoretical limitations of the studies presented alongside ways that these general findings may be built upon by further applied and theoretical research.

Chapter 1: General introduction

1.1. Introduction

Human beings are a remarkable species. We have achieved great advances in technology and science, and successfully colonised almost every environment on earth. This, really, is made possible through the combined effects of two behavioural traits, cooperation and social learning. The extent that humans cooperate and use social learning is unparalleled in the animal kingdom, allowing humans to overcome extreme environmental difficulties, develop continually improving solutions to problems and to live (relatively) peacefully alongside one another. Though animals do engage in cooperative behaviour (Clutton-Brock, 2009; Dugatkin, 2002) and also use social learning or form behavioural traditions akin to culture (J. R. Kendal et al., 2009; Perry, 2011; Vale, Flynn, et al., 2017), this is not to the extent displayed by humans. As such, understanding the evolution and manifestation of cooperation and social learning in human beings provides a vital insight into understanding the evolution of human societies and the reasons why our species is so successful.

The broad overall aim of this thesis is to evaluate the extent to which cooperative behaviour might be influenced by social learning or the influence of other individuals more broadly. The potential importance of social learning to explaining human cooperative behaviour has been theorised within cultural evolution (Boyd & Richerson, 2009a; Henrich, 2004; Richerson et al., 2016; D. Smith, 2020) but there is a lack of direct empirical testing. As will be alluded to in the proceeding sections and throughout the rest of the thesis, cooperative behaviour comes in many forms and is expressed in a variety of ways. As such, this question is addressed by using a variety of approaches, including experimental, field, quantitative and qualitative approaches (though, the focus is primarily quantitative). The overarching framework of the thesis is evolutionary and draws heavily on insights from the cultural evolution literature. I will therefore investigate evidence for the influence of social learning strategies on cooperative, selfish and spiteful behaviour in a traditional experimental setting but also explore this in an applied and more ecologically valid setting. I will also use a simulation model to investigate what the likely consequence of the experimental cooperation I observe would be over a longer timeframe and alongside selection. Though the empirical focus of this thesis is to investigate this question specifically in humans, throughout, and where relevant, I also draw upon insights from studies into non-human animals.

1.2. Cooperation

Cooperation, defined henceforth as behaviour which is costly or beneficial to the actor and benefits another individual(s) (and is selected for based on its cooperative merit[1]), (West et al., 2007b), is a puzzle that has fascinated evolutionary theorists since Darwin's original writings. The puzzle is such that, although mutual cooperation presents the best collective outcome for both parties, an individual who abstains from this effort but still receives the benefit would be individually better off. Because natural selection predicts that individuals should invariably behave in ways that maximise their own personal fitness (Darwin, 1859), it would predict the evolution of indiscriminately selfish individuals. Such situations, where one's personal interests do not line up with the interests of the group, have been termed social dilemmas (Dawes, 1980; Kollock, 1998; Van Lange et al., 2013).

The most famous formulation of a social dilemma is the prisoner's dilemma (Rapoport, 1974) which is described as follows: two individuals are detained by the police and placed into separate rooms and there is sufficient evidence to convict both individuals of a minor crime, but they are suspected of a larger crime. The police independently offer both parties the same deal of testifying against their partner in exchange for being released without charge. The dilemma is that, if one individual remains silent but their partner does not, the individual who remains silent will be charged with the major crime while the other individual escapes without punishment. Therefore, regardless of the decision of your partner, the rational decision is to defect and testify. This protects you from exploitation if your partner testifies but presents you the chance to obtain the best personal outcome in the event that your partner remains silent.

[1] The additional caveat on this definition is to exclude otherwise selfish actions which, by chance, happen to benefit another individual. For example, although the dung beetle benefits from the production of elephant dung, this would not be considered a cooperative act on the part of the elephant.

This dilemma can easily be expanded to consider any number of individuals (sometimes referred to as an N person prisoner's dilemma), originally described by (Hardin, 1968) as the "tragedy of the commons". Suppose there is a group of farmers, each with several cattle, and a public moorland (the "public good") which provides space for all the farmers to rear their cattle. Further suppose that this moorland cannot support all the cattle feeding at once or else it would deteriorate and leave no land for anyone. In this scenario, some individuals must behave selflessly and refrain from grazing their cattle. However, whoever does so, will lose out on the benefits of the moorland while everyone else benefits (or, free rides) from their selfless act. Consequently, in the absence of other reasons that farmers might refrain from grazing their cattle[2], every farmer will maximise their own personal benefit and the moorland will disappear. This example highlights that cooperative behaviour need not actively benefit another individual, it may instead take the form of simply refraining from engaging in selfish behaviour (J. Heath, 2006).

These dilemmas may seem somewhat abstract, but applications of them continue to be relevant for important global issues facing humanity today (Lange & Waring, 2018). The climate can be thought of as a public good from which the whole world benefits. Globally, all countries (to varying extents) contribute to the emission of CO_2 which has consequences for the rest of the world. However, reducing emissions by rebuilding infrastructure is costly which means there is a strong incentive to free ride on the efforts of countries that do make the investment. Other smaller scale examples of cooperative dilemmas include voting (where one's individual vote makes little difference but maintains the integrity of the democratic system), overfishing, harvesting trees for palm oil or the choice to drive rather than walk or use public transportation.

Nevertheless, human beings cooperate to a degree that is, for the most part, unequalled by non-human animals (perhaps, precisely because we are humans: (Rothschild, 2009). Where animal cooperation is often based on kinship (J. E. Smith, 2014) or in some limited cases, reciprocity (Carter & Wilkinson,

[2] Chapter 2 will discuss in detail the various reasons why an individual may behave altruistically, despite the risk of exploitation or the benefits of behaving selfishly.

2013) or punishment (Truskanov et al., 2020), humans remain the only known species to engage in large scale, often unreciprocated cooperation with unrelated individuals (Fehr & Rockenbach, 2004). For example, £32.8m in donations were made to the NHS in support of Sir Captain Tom's endeavours during the COVID-19 pandemic (BBC News, 2021), and extreme examples of altruism were shown towards fellow victims of the London tube bombings in 2005 (BBC News, 2011). Ethnographically, human societies are characterised by their proclivity towards cooperative behaviour (Hill et al., 2014; Kurzban et al., 2015) and this tendency has been suggested to be a large part of the reason why human societies are so generally successful (Smaldino et al., 2013). Cooperative behaviour may also have driven the transition from medium to larger scale societies and explain how we have been able to colonise the earth so widely[3] (Powers et al., 2016) as sustained cooperative efforts allow individuals to achieve things impossible alone (Heath, 2006). The benefits and importance of humans' cooperative tendencies cannot be understated.

1.3. Social learning

Humans are a cultural species; across the globe there are a myriad of different social traditions, languages and counting systems which have been developed and transmitted between individuals for generations. Though it is generally agreed that humans have culture, there are a huge variety of ways that culture has been defined (Ramsey, 2013). As a working definition, it may be defined as "group-typical behaviour patterns shared by members of a community that rely on socially learned and transmitted information" (Laland & Hoppitt, 2003, p. 151). Animals also possess culture to a certain extent (Laland & Hoppitt, 2003; Ramsey, 2013; Whiten, 2019b). Primates (Perry, 2011), birds (Aplin, 2019) and fish (Laland & Williams, 1997) have all been shown to develop stable and socially transmittable behavioural traits, which many argue provides evidence of culture.

[3] Though, it has been suggested that population size itself may influence the evolution of cooperation such that reciprocation is only favoured in larger populations (Hilbe et al., 2018).

What animals lack, or at least show very little evidence of, is a more powerful form of culture, termed cumulative culture. Like culture, there is not a universally agreed-upon definition of cumulative culture[4], but definitions, in general, agree that, quintessentially, cumulative culture involves the transmission of traits across generations with some small modification or improvement (Mesoudi & Thornton, 2018; Vale et al., 2017). For example, building aeroplanes did not arise from a single individual, it was the result of generations of iterative improvements. Currently, there is little evidence of cumulative, "ratchet-like", culture in any animal species (Dean et al., 2014; Mesoudi & Thornton, 2018), but it is one of the key reasons why human beings are so successful (Boyd et al., 2011; Legare, 2017).

For culture to exist at all, however, individuals must be able to acquire information from one another and for said information to result in persistent changes in behaviour[5]. This process is termed, social learning, defined as learning facilitated by observing or interacting with another individual or their products (Heyes, 1994). Social learning has long been considered in the social psychology literature (Bandura, 1986) but has gathered specific attention within the cultural evolution field. Social learning provides individuals the opportunity to exploit social information about their environment without risking the potential costs associated with learning for themselves. For instance, copying food sources from your conspecifics might be an effective way to avoid food poisoning. Instances of social learning have been found in a variety of animal species including birds (Aplin, 2019), fish (Laland & Williams, 1997) and various other vertebrate and invertebrate species (Whiten, 2019b). However, copying also introduces the risk of acquiring outdated or harmful information (Giraldeau et al., 2002; Mesoudi, 2009; Mesoudi & Lycett, 2009; Rieucau & Giraldeau, 2011). Therefore, some degree of asocial learning or innovation must persist alongside social learning to produce solutions to novel

---

[4] To provide an idea of the range of definitions. Some focus on the production of behaviour unattainable by a single individual (Boyd & Richerson, 2005), while others focus on increases in complexity and/or efficiency (Dean et al., 2014). The classic description describes behaviour ratcheting up in complexity or efficiency across generations (Tomasello et al., 1993).

[5] Recently, authors have argued that many more mechanisms (besides just social learning) go into cultural evolution (for example, neural and contextual factors) and warn against restricting focus to social learning strategies (Singh et al., 2021).

environmental challenges (R. Kendal et al., 2005; Legare & Nielsen, 2015; Schlag, 1996a). Furthermore, the fitness gains associated with social learning in a variable environment decreases with the frequency of those using it (Rogers, 1988). For these reasons, unbiased or random copying is unlikely to be favoured by natural selection (Boyd & Richerson, 1995). The solution is the adaptive use of "social learning strategies" or transmission biases which influence when, what, and from whom individuals socially learn (R. Kendal et al., 2018; Laland, 2004; Morgan et al., 2012). These social learning strategies have been the focus of sustained theoretical and empirical interest in recent decades. Particularly, applying cultural evolution to the study of sustainability represents a growing area of interest (Brooks et al., 2018). However, little is known about the potential role of social learning, and social learning strategies, in explaining the widespread persistence of human cooperation, though this too represents a growing area of interest. In this thesis I aim to address this question by applying several different methods to evaluate the influence of social learning on cooperative, selfish and spiteful behaviour.

1.4. Aim and structure of the thesis

In this thesis, I seek to build upon the cooperation and social learning literature by investigating underexplored topics and themes within the field of cooperation and expanding the still relatively novel field of research applying social learning to cooperative behaviour. In chapter 2, I review the literature surrounding cooperation, social learning and the combination of the two in primarily (but not exclusively) humans. In the first empirical chapter (chapter 4), I use an online experiment to assess the explanatory role of three social learning strategies in explaining participant cooperation in a Prisoner's Dilemma and a Snowdrift game. In chapter 5 I assess success based and conformist transmission in explaining spiteful or antisocial behaviour in an online experiment. In chapter 6, I present the results of a questionnaire and follow up interviews exploring Durham University students' attitudes towards environmental sustainability, with particular focus on the role of social learning or social influence more broadly. Additionally, the questionnaire also includes a behavioural measure, to allow a further experimental comparison of the influence of the three social learning strategies in an applied context. Finally, in chapter 7, I conclude the thesis with a general discussion.

Each of the three empirical chapters (chapters 4,5, and 6) are written in the form of papers for publication. Consequently, there is some overlap between the introductory material presented in chapter 2 and the introduction sections of the empirical chapters. While all the research was conducted by RW, each of the papers will be multi-authored, reflecting the contribution of supervisors and an external collaborator, who will also have authorship on chapters 4 and 5. As such, the empirical chapters refer to "we" and "our" rather than "I" throughout. At the time of writing, chapter 4 has been published here:

Watson, R., Morgan, T. J. H., Kendal, R. L., Van de Vyver, J., & Kendal, J. (2021). Social Learning Strategies and Cooperative Behaviour: Evidence of Payoff Bias, but Not Prestige or Conformity, in a Social Dilemma Game. Games, 12(4), 89. MDPI AG. Retrieved from http://dx.doi.org/10.3390/g12040089

Following COVID-19 restrictions, the design of chapter 6 was changed from the original plans, however the overarching research question (the influence of social learning on sustainable behaviour) remained the same.

Chapter 2: Literature review: The interaction of cooperation and social learning

2.1. Introduction

In this chapter, I review the cooperation and social learning fields that are drawn upon in this thesis. Because of their size, it is not possible to provide a wholly inclusive review of every corner of the respective fields. Instead, I provide an overview of the most important findings and concepts. The first section highlights the major explanatory theories offered to account for the evolution of cooperation and, whenever possible, describes how they have been addressed experimentally. I take a similar approach for social learning and discuss content or context based social learning strategies. The final section considers work that has brought the two fields together. This includes a brief discussion of cultural group selection (Richerson et al., 2016), a major theory as to how social learning (and cultural evolution) could account for the unusually widespread cooperative behaviour in human societies and discussions on how cooperation and social learning has been applied to societal issues. As a note of caution, most of the research discussed in this section is restricted to WEIRD (western educated industrialised rich and democratic) populations, which may not be representative of universal human trends (Henrich et al., 2010). The relatively little research addressing cross cultural differences in cooperation is discussed in section 2.2.3.6. I then conclude by considering how theories from cooperation and cultural evolution can be applied to the problem of sustainability (section 2.5).

2.2. Cooperation

The working definition of cooperation I apply in this thesis is, behaviour that benefits a second party at either a cost or benefit to the actor (West et al., 2007b). This definition captures altruism and some mutualistic actions but excludes selfish behaviour with an unintended positive by-product (see section 1.2). Explanatory factors are also often split into two different groups, depending on what kind of fitness they act upon (see figure 2.1). Direct fitness is measured by the impact of an individual's behaviour on the production of their own offspring while indirect fitness is the gain in fitness associated with the reproduction of an individual's relatives (West et al., 2007b). Both kinds have been vitally important for the evolution of cooperation and will be considered in the proceeding section.

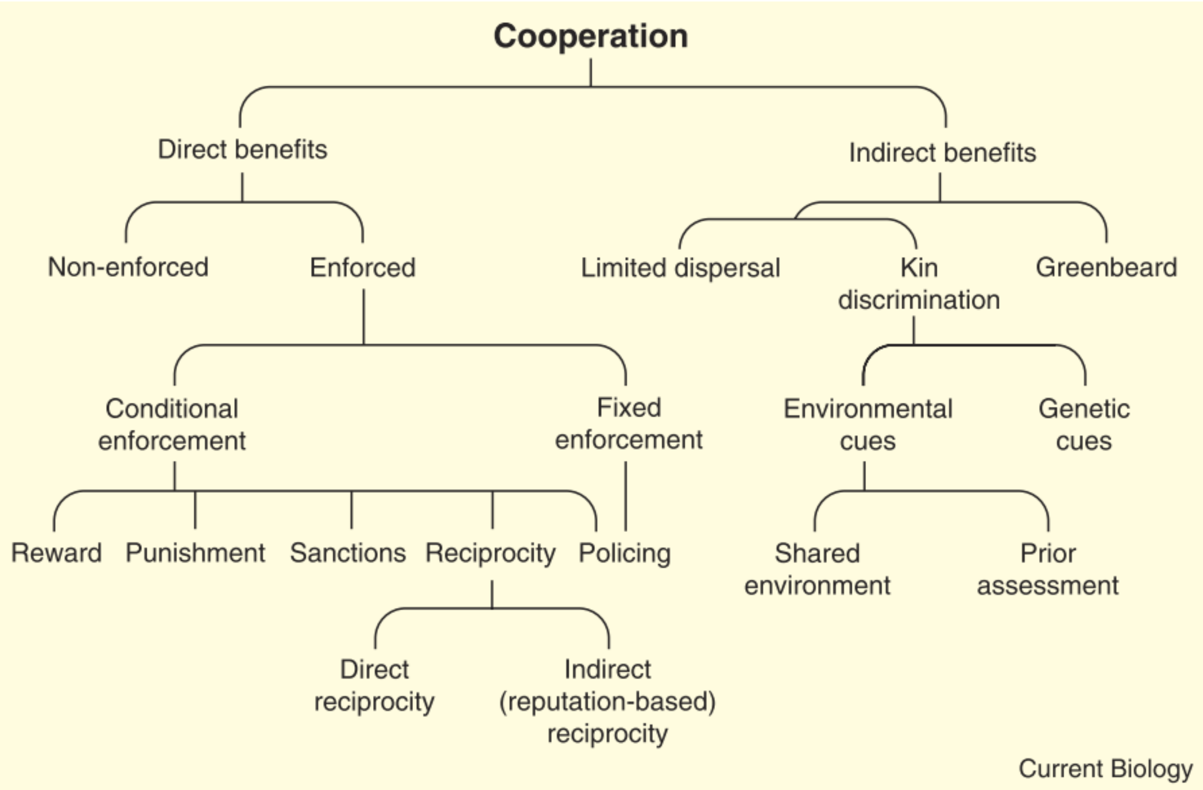


Figure 2.1. *Diagram taken from* West et al., (2007a) *showing the distinction between direct and indirect benefits. Most (but not all) of these factors are discussed in this chapter.*

2.2.1. Indirect benefits

2.2.1.1. Inclusive fitness

Inclusive fitness (or, kin selection) is one of the oldest explanatory accounts for cooperative behaviour. Proposed by William Hamilton in his now seminal paper "*the genetical evolution of social behaviour*" (Hamilton, 1964), he suggested that cooperation may be genetically favourable provided it is directed towards genetically related individuals. Therefore, though cooperation may be costly at the phenotypic level, it may be altruistic at the genetic level. This can be summarised by a simple equation. If the benefits *(b)* to another individual multiplied by the coefficient of relatedness *(r)* minus the costs *(C)* of the behaviour are greater than 0, then altruism can be favoured through inclusive fitness.

$$r\,b - C > 0$$

Inclusive fitness has been used to explain a huge variety of cooperative behaviours, especially in the case of animals (Griffin & West, 2002; J. E. Smith, 2014). For example, brush tailed possums preferentially share dens with close kin when den availability is scarce (Banks et al., 2011) and female spotted hyenas were more likely to assist close kin in aggressive encounters than less related individuals (J. E. Smith et al., 2010). In coalitionary matings in wild turkeys, it was possible to directly calculate the components of Hamilton's equation which showed the reproductive gains of the dominant individual outweighed the reproductive costs of the genetically related helper individual (Krakauer, 2005). Inclusive fitness has also been used to help explain eusociality in social insects, though this has been disputed (Nowak et al., 2010).

In humans, though anthropologists have long identified kinship as an important determinant of cooperation (Kurland & Gaulin, 2005), fewer studies focus on inclusive fitness in humans than in animals. This may be because kin selection is viewed as a given part of human societies and therefore uninteresting (Rand & Nowak, 2013) or because unreciprocated non-kin cooperation is what makes human cooperation unique (Fehr & Rockenbach, 2004). Nevertheless, there are some examples of kin effects on cooperative behaviour in humans. One experiment showed individuals would perform painful wall sit exercises for longer to earn monetary rewards for close family than their friends or a charity (Madsen et al., 2007) and social network analysis showed cooperation between Chinese farmers was more likely between close kin (Thomas et al., 2018).

2.2.1.2. Green beards

Green beard cooperation is a similar concept to kinship and was proposed as a possible kin recognition mechanism (Hamilton, 1964). In short, individuals may possess alleles which produce a recognisable trait and bias cooperation towards others with that trait. The term came later from Richard Dawkins (1976) who imagined a population of individuals with green beards who preferentially helped other green bearded individuals. There are two problems with this theory that mean green beard genes are unlikely to be common. Firstly, it is hard to imagine a single gene simultaneously coding for both cooperating with, recognising, and signalling to fellow green beards (Davies et al., 2012). Secondly, it would be relatively easy for false beards (display the beard but do not cooperate) to invade a population of green beards (Gardner & West, 2010). Nevertheless, examples of green beard cooperation have been documented.

Strands of slime mould that carry the cSA gene selectively provided supporting stalks to other carriers of the gene in times of food shortage (Queller et al., 2003). Similarly, in yeast cells, carriers of the costly (in terms of growth rate) $FLO1^{+}$ gene selectively cluster with other carriers of the gene. This clustering allowed the yeast to resist a larger variety of environmental stressors (Smukalla et al., 2008). Fire ants show an extreme form of green beard cooperation as workers execute queens who carry a BB allele or other workers that have been rubbed with the scent of BB carrying queens (Keller & Ross, 1998).

There is little evidence for green beard cooperation in vertebrates and especially humans. Though some have suggested that language or accent may act as a green beard trait in humans (E. Cohen, 2012). However, there is little reason to assume that the genes coding for accent are related to those coding for cooperation (Gardner & West, 2010). Furthermore, simulation work in well-mixed populations found the false beard problem results in the continuous cycling of traits that exploit established cooperation (García et al., 2014).

2.2.1.3. Spite

Besides helping relatives, inclusive fitness provides one additional way for individuals to assist their kin. Spiteful behaviour, defined as any behaviour which is costly to both the actor and recipient (West et al., 2007b), could be used to eliminate competition for one's own kin. This kind of spite was later termed *evolutionary spite* (Jensen, 2010). Both Hamilton (1964) and Wilson (1975) suggested that spiteful behaviour could be favoured if it was directed at individuals who were less related to the actor than a random member of the population but concluded that spite is unlikely to be generally important for social behaviour. Convincing examples of spite are extraordinarily rare, as it is difficult to demonstrate that a behaviour costly in the short term would not result in tangible direct benefits, to the actor, somewhere within their lifetime (Foster et al., 2001). It is also hard to imagine many scenarios where the best strategy available to enhance your reproductive success is to harm others at a cost to yourself (West & Gardner, 2010). Only under intense competition is this likely to be the case (Gardner & West, 2004; Jensen, 2010). The most convincing example of evolutionary spite are parasitic wasps which sometimes hatch into sterile soldiers (thus, sacrificing all reproductive ability) and attack less related individuals (Gardner et al., 2007). Two other examples are honeybees which destroy unfertilised worker bee eggs to protect the reproductive monopoly of the queen (Barron et al., 2001) and red fire ants which execute queens with BB allele (Keller & Ross, 1998). In both the honeybee and fire ant case, besides the energetic costs, the behaviour is not costly and so is not consistent with the strict definition of spite (West et al., 2007b). There are very few examples of evolutionary spite in non-human vertebrates, though humans do often display spite in experimental settings (Abbink & Sadrieh, 2009). The topic of spite will be returned to in chapter 5.

2.2.2. Direct benefits

The following section details explanations and factors that change or influence the personal costs and benefits of cooperation. This can be done through selectivity in one's choice of partner, or by encouraging a partner's cooperation through other means. These factors help explain cooperation beyond relatives which is of particular interest for human cooperation which frequently occurs in large unrelated groups (Kurzban et al., 2015).

2.2.2.1. Direct reciprocity

Under a prisoner's dilemma (PD), game theory predicts that defection is always favoured (Rapoport, 1974). However, it is rarely the case that individuals meet in only a single interaction. Over repeated interactions, individuals could reciprocate cooperation quid pro quo[6], which reaps the benefits of cooperation without the risk of exploitation. This has been termed, direct reciprocity (Hilbe et al., 2018; Trivers, 1971). In a computer tournament playing repeated PD games, the strategy Tit for Tat (match partners decision) outcompeted all other strategies (Axelrod & Hamilton, 1981). Although future work identified strategies which outcompete TFT (Nowak & Sigmund, 1993), the fundamental tenant of direct reciprocity remains. Cooperation can be favoured through direct reciprocity where the probability of a future interaction (w) outweighs the costs (c) and benefits of the behaviour (Nowak, 2006).

$$w > c/b$$

In experiments, humans display a remarkable willingness to reciprocate cooperation received from others (Gintis, 2000; Gintis et al., 2003). In repeated games, individuals in fixed groups establish markedly higher rates of cooperation than varying groups where reciprocal relationships cannot be established (Duffy & Ochs, 2009; Fehr & Schurtenberger, 2018). Non-cooperation is often accompanied by guilt (Ketelaar & Tung Au, 2003) and participants generally believe cooperation to be the morally appropriate behaviour, unless others fail to cooperate, indicating a concern for reciprocity (Kimbrough & Vostroknutov, 2016). Children aged 3 months selectively play and share

[6] The Latin phrase meaning "a favour or advantage granted in return for something". Colloquially, I scratch your back you scratch mine.

with those who have previously displayed prosocial behaviour (Leimgruber, 2018), suggesting such norms are understood early in development

Contrary to classic assertions (West et al., 2007a), examples of direct reciprocity have also been found in animals. Vampire bats share food with unrelated nest mates in a manner not explained through harassment or errors in kin recognition (Carter & Wilkinson, 2013). Grooming in Barbary macaques is directed towards individuals who have previously offered grooming (Molesti & Majolo, 2017). Rats have also shown a willingness to reciprocate cooperation in iterated prisoner's dilemma games (Wood et al., 2016) and to extend reciprocal altruism in one task to a different unrelated task (Schweinfurth & Taborsky, 2017).

While reciprocity may be a powerful proximate cue in cooperative decision making, it has been argued that it cannot account for how cooperation would initially evolve in a population of defectors (for example, Boyd & Richerson, 1988). More recent models have challenged this view, showing that under a stochastic (rather than deterministic) process, a single TFT player can invade a population of defectors (Nowak et al., 2004). Direct reciprocity is also more stable if shares of a public good do not decrease with increasing group size (Kurokawa & Ihara, 2017) or if individuals behave strategically and cooperate when it is beneficial (Rothschild, 2009). A further condition for direct reciprocity is individuals must be forgiving towards occasional defection (as in, generous TFT (Nowak & Sigmund, 1992)), otherwise populations become locked in cyclic defection.

2.2.2.2. Indirect reciprocity and reputation

While direct reciprocity proposes a viable mechanism by which individuals can develop reciprocated cooperative relationships, it cannot explain altruistic behaviours like charity donations that have little scope for reciprocation. Individuals will also encounter potential cooperative partners about whom they have no prior information. The theory of indirect reciprocity addresses this by suggesting that individuals should bias cooperation towards those who have behaved cooperatively to others (Nowak & Sigmund, 1998). In other words, individuals should consider reputations and choose partners accordingly. Like direct reciprocity, indirect reciprocity requires that co-operators are forgiving towards occasional defections (Ohtsuki & Iwasa, 2006) but also that individuals only receive a bad reputation for failing to cooperate with co-operators (Nowak & Sigmund, 2005; Ohtsuki & Iwasa, 2004). However, indirect reciprocity may be less stable than direct reciprocity if individuals have a high chance of repeated interactions (Roberts, 2008). It is likely that direct and indirect reciprocity complement one another to maintain cooperation in a population. Another related possibility is that a cooperative reputation might serve as a costly signal of mate quality, though overall evidence for this is mixed (Bhogal et al., 2019; Roberts, 1998).

General reputational effects have been shown in humans in several experiments (for a review see; Wu et al., 2016). Participants increased their cooperation towards individuals that publicly donated to a climate change fund (Milinski et al., 2006) and were motivated to increase their cooperation if the two highest or lowest donations were publicly highlighted (Jacquet et al., 2011). However, though individuals in a cooperative game appeared to base their decisions on third party information about their partner's behaviour, they ignored information about the behaviour of their partner's last partner (Milinski et al., 2001). In field contexts, posters of eyes (indicating a cue of being watched) have been shown to encourage cooperation in a variety of contexts (Bateson et al., 2006; Dear et al., 2019; but see Rotella et al., 2021). In a car park, drivers were more likely to give way to an oncoming car if the previous car they met had given way to them (Mujcic & Leibbrandt, 2018).

Examples of indirect reciprocity are far less common in animals, which may be because of the increased cognitive demands associated with indirect reciprocity (Nowak & Sigmund, 2005). There are, however, some limited examples. An experiment showed that after simulating an intrusion by A into a neighbouring (B) song sparrow's territory (by playing As birdsong) other song sparrows then ceased to restrain their aggression towards A, which indicated the use of third party reputation information (Akçay, Çağlar et al., 2010) and clients of cleaner fish spent more time next to cleaners they observed to be cooperative with a third party (Bshary & Grutter, 2006). Dogs however were found to be completely insensitive to cues of eyes to prevent the stealing of food (Neilands et al., 2020).

2.2.2.3. *Network reciprocity*

Many evolutionary models assume populations are well mixed, such that every individual has a chance of interacting at some point (Nowak, 2006). This is not consistent with how real populations are structured where individuals freely choose whom they assort and interact with (Nowak, 2006). An early simulation study identified that spatial assortment alone can be enough for cooperation to evolve (Nowak & May, 1993). Since then, many studies have investigated various kinds of spatial structure and the evolution of cooperation is often highly sensitive to these factors. In general, cooperation can evolve through selective assortment, provided the benefit of the cooperation (b) divided by the cost (c) is greater than the average number of neighbours (k) (Ohtsuki et al., 2006).

$$\frac{b}{c} > k$$

Experimental results for the impact of spatial structure have been mixed. One experiment comparing fixed and random groups varied the benefits of cooperation and the number of neighbours and found that cooperation was stable when the above equation was satisfied (Rand et al., 2014). Conversely, other experiments have found cooperation rates were equivalent across varying population structures (Gracia-Lazaro et al., 2012). A more recent experiment showed that well mixed groups, provided some ties between individuals varied between rounds, was sufficient to maintain cooperation between those with unchanging ties (Harrell et al., 2018). For further reviews and discussions of the effect of spatial structure on cooperation, see Perc and Szolnoki, (2010) and Roca et al., (2009).

2.2.2.4. *Punishment*

Sometimes it may be necessary to force individuals to cooperate. A substantial amount of work has focused on exploring the role of costly punishment in maintaining cooperation. Elinor Ostrom (1990) argued the importance of institutions for solving social dilemmas and that such institutions should be expected to emerge naturally through cultural evolution. Early models supported the beneficial role of punishment, showing that peer punishment can effectively maintain cooperation in large groups, provided punished individuals switch to cooperation (Boyd et al., 2010; Boyd & Richerson, 1992). One issue with peer punishment is second order free riding, where co-operators who do not punish outcompete co-operators who do punish. There are several solutions to this problem, including the punishment of non-punishers (Boyd & Richerson, 1992), pooling the costs of punishment across multiple individuals (Sigmund et al., 2010) or the formation of dedicated institutions that punish non-co-operators (Ostrom, 1990). Additionally, punishment could also spread between individuals via social learning (Andrés Guzmán et al., 2007; FeldmanHall et al., 2018; Li et al., 2021).

Although there is a theoretical incentive to refrain from punishing defectors (as it is costly and provides no direct benefit), a wealth of experimental evidence has shown that participants are willing to punish non-cooperators (Balliet, Mulder, et al., 2011; Gintis, 2000). Individuals frequently reject (and thus punish) low offers in ultimatum games (Güth & Kocher, 2014) or pay a cost to reduce the earnings of low contributors in public goods games (Fehr & Gachter, 2000). Punishers are sometimes rewarded by their group mates (dos Santos et al., 2013) but the proximate reasons for punishment are not clear. Neurological evidence in humans has shown activity in regions of the brain associated with pleasure when punishing (De Quervain et al., 2004), and other evidence points towards a desire to enforce cooperative norms (Carpenter et al., 2004). However other experiments have shown participants reported envy, not anger, when individuals earned more than them (Pedersen et al., 2013) and when given the choice would rather compensate those who missed out than punish defectors (Pedersen et al., 2018).

There are several other factors worth attention with respect to punishment. There is considerable variation across cultures in the extent (and kind) of punishment that is used (Henrich et al., 2001; Herrmann et al., 2008; Marlowe et al., 2008). Relatedly, in experiments, defectors will sometimes direct punishment towards co-operators, which in turn prevents punishment from encouraging cooperation (Rand & Nowak, 2011). Punishment is also rarely seen in animals (aside from some limited examples (Raihani et al., 2012)) and usually only occurs when power is unequal or there are other direct benefits to the behaviour. Therefore, punishment is likely a key influence in human cooperation only.

2.2.3. Other factors

Every mechanism discussed so far accounts for cooperation despite it being, ultimately, maladaptive (it is costly to help other individuals). Most of the cooperation research only considers cooperative scenarios that are consistent with a PD, perhaps because this is where it is most difficult to explain. There are other scenarios where cooperation may not always be maladaptive. For example, there are instances where the optimal behaviour might depend on what your partner is doing. Two such payoff structures I will discuss in this section are the Snowdrift (SD) game and the Stag hunt game (SHG). In other instances, the benefits associated with cooperation are large enough that they outweigh the possible gains from defection, such that cooperative behaviour can be mutualistic. Sections 2.2.3.1-3 covers these alternative scenarios. Sections 2.2.3.4-6 then describes three additional factors that can also support the maintenance of cooperative behaviour.

2.2.3.1. Snowdrift games

The most well-known alternative to the prisoner's dilemma is the snowdrift (SD) game (or chicken / hawk dove game)[7](Doebeli & Hauert, 2005a). In this game, two drivers meet on a road blocked by snow, which must be cleared for the drivers to continue. Although cooperating would clear the snow quickest, if one driver gets out to begin shovelling, it may be in the interest of the other driver to remain warm in their car. Consider the payoff matrices below (Table 2.1). Unlike the PD, in the SD game, defection is only favourable if your partner is cooperating. Otherwise, cooperation is the favourable alternative. The SD game applies to scenarios where some mutual benefit is unobtainable without a minimum investment, such as flood or predator defences.

---

[7] These terms are often used synonymously though the payoff matrices do sometimes differ. For the purpose of this thesis, I am assuming that these terms refer to the same payoff matrix.

Table 2.1. *Payoffs associated with cooperation and defection) depending on the behaviour of a partner between a PD and SD game adapted from* Doebeli and Hauert (2005). *Both tables show benefit (b) compared to the costs (c). Payoff rankings for the column individual are DC > CC > DD > CD for the PD game and DC > CC > CD > DD for the SD game.*

	Prisoner's dilemma (PD)		Snowdrift (SD)	
	Cooperation	Defection	Cooperation	Defection
Payoff to Cooperator	*b – c*	*-c*	*b – c/2*	*b - c*
Payoff to Defector	*b*	0	*b*	0

One of the earliest applications for the SD game was to understand why animals rarely used weapons to inflict serious damage. The solution lay in understanding that while individuals fighting to do damage (hawks) are favoured against individuals who never escalate conflicts (doves), whereas at the population level, doves are favoured against hawks as they avoid the costs of losing in conflicts. Consequently, the evolutionarily stable strategy (ESS) between hawks and doves is a stable population of both strategies (Maynard Smith & Price, 1973). This dynamic is predicted between cooperation and defection in the SD game, where the ESS is a population comprised of both strategies (Doebeli et al., 2006; Doebeli & Hauert, 2005a; Kun et al., 2006; Souza et al., 2009), which the authors dubbed the "tragedy of the commune" (Doebeli et al., 2006). Yeast cells also showed this pattern, where they formed stable populations of cells which secrete costly mutually beneficial enzymes and cells that don't (Gore et al., 2009). Another difference between the SD and PD games is while spatial structure facilitates the formation of localised networks of co-operators in the prisoner's dilemma (Nowak & May, 1993) this does not occur in the snowdrift game. Instead, dendritic spines of co-operators form (Hauert & Doebeli, 2004).

There are considerably fewer experimental studies investigating the SD game than the PD and those that do use the SD game often restrict cooperation and defection to binary choices (rather than continuous, as in typical public goods games). General findings point towards higher levels of cooperation in the SD than the PD (Hilbig et al., 2018; Kümmerli et al., 2007; Reed et al., 2018; Su et al., 2018; Wit & Wilke, 1992). A few snowdrift models do consider cooperation on a continuum and generally find that cooperation converges towards investments of around 50% per round (Brown & Vincent, 2008; Mcnamara et al., 2008; Sasaki & Okada, 2015; Zhong et al., 2008). To my knowledge, no experiments have considered continuous SD games over multiple rounds which remains a key empirical gap. SD games are discussed further in chapter 4.

2.2.3.2. Stag hunt games

A further alternative to the PD and SD is the stag hunt game (SHG) (Skyrms, 2004). In the SHG, a group of hunters are lying in wait to capture a stag, which will be successful, provided nobody gives away their position. Successfully capturing the stag represents the best outcome for all hunters. But the hunters are not certain that the stag will come. Before the stag arrives, a hare is seen moving along the path where they have set their trap. If a single hunter jumps out and captures the hare, they will eat but this will waste the trap set for the stag and all the other hunters will go hungry. The payoff matrix for the SHG is shown in Table 2.2. The payoff matrix is such that if your partner is cooperating, you do best by cooperating, whereas if your partner is defecting, you do best by defecting. In other words, this introduces a dynamic of trust (Bolle & Spiller, 2021; Skyrms, 2004). One consequence of this is SHG are more likely than PD games to produce dynamics of in group bias where knowing and understanding the intentions of your partner is important (Jansson, 2015).

The SHG results in two possible population states, full cooperation or full defection (Bolle & Spiller, 2021). Within experiments comparing binary PD, SD and SHGs, participants displayed generally higher cooperation in the SHG (Hilbig et al., 2018). Other experiments have explored the proximate reasons for participant’s decisions. In two player SHGs, cooperation was more likely when an individual was more trusting and also when their partner was described as trusting (Jansson & Eriksson, 2015). Büyükboyaci (2014) showed that cooperation in the SHG was less likely when individuals perceived that their partner was risk averse but was unrelated to a participants own risk aversion. Comparing the PD with the SHG, while PD cooperation was driven by beliefs that it is the morally right behaviour, SHG cooperation was driven by beliefs of the efficiency of cooperation

(Capraro et al., 2020). Due to practical and financial limitations, the SHG is not considered further in this thesis.

Table 2.2. *Payoffs associated with cooperation and defection for the left (row) player depending on their partners decision (column) in the stag hunt game. The payoff hierarchy is CC > DC > DD > CD.*

	Cooperation (C)	Defection (D)
Payoff to cooperator (C)	4	1
Payoff to defector (D)	3	2

2.2.3.3. Mutualism

Some scenarios are such that the payoffs associated with cooperation are so large that they outweigh any potential gains from free riding on cooperators. These scenarios have been informally termed the Prisoner's delight (Binmore, 2007a), but generally represent cases where there is no social dilemma. There are several ways that this may be realised.

Interactions between individuals can be mutualistic, where individuals provide simultaneous (but personally costly) benefits to one another (Leigh, 2010). This is especially common in microbes and bacteria (Frederickson, 2017). For example, rhizobium bacteria provide N_2 to the host plant in exchange for tolerance and O_2 from the host (Kiers et al., 2003). A different example is the parasite removal service provided by cleaner fish where both parties benefit (Bshary & Grutter, 2006). A similar scenario arises when individuals produce a benefit to others as a by-product of otherwise selfish behaviour (though, this often does not qualify as cooperative behaviour per se (West et al., 2007b)). Ant foundresses tolerate one another when establishing the nest (enjoying the direct benefits of faster nest building) but then fight to monopolise reproduction once the first brood of workers emerge (Bernasconi & Strassmann, 1999).

An individual's fitness may also be directly tied to another individual. In this case, there is no temptation to defect because the payoff maximising strategy is to cooperate with your neighbour (Roberts, 2005). In termites, two queens (but no more than two) will produce a brood cooperatively when a male is not available, with the larger queen providing additional resources to help the smaller queen grow which increases the colonies' resistance to parasites and its overall growth rate (Matsuura et al., 2002). Another example is cooperative hunting of red colobus shown by chimpanzees (Boesch & Boesch, 1989), though some argue this is an example of a selfish by product rather than cooperation (Tomasello et al., 2012). For humans, it has been suggested that the cognitive mechanisms required for interdependent cooperative hunting developed into those required to maintain cooperative norms (Tomasello et al., 2012). Consistent with this, young children, unlike chimpanzees, engage with, and understand the commitments of, joint activities (Gräfenhain et al., 2009). Interdependence may be of greater importance in explaining human cooperation than other animals. Finally, individuals may be incentivised to engage in cooperative breeding because of the benefits of living in a larger group, termed group augmentation (Kingma et al., 2014). Meerkats are especially well known for this, where breeding is restricted to the dominant pair, but childcare is provided by many individuals within the group (Clutton-Brock et al., 2001).

2.2.3.4. Religiosity

Religiosity has been proposed as a mechanism to maintain cooperation. A common feature of many religions is a moralising and omniscient supernatural agent who punishes bad behaviour (Norenzayan & Shariff, 2008). Bulbulia (2004) proposed a revised prisoner's dilemma which included the existential (divine) consequences of choosing defection, which then incentivises cooperation out of fear. This would also suggest that individuals should bias their cooperation towards members of their religious group to avoid the risk of free riding (Bulbulia & Sosis, 2011). A review of 19th century US communities found that religious communities with the costliest joining requirements (and thus more difficult to fake membership) survived the longest, implying they more successfully solved collective action problems (Sosis & Bressler, 2007). Costly requirements did not, however, predict the survival of secular societies (Sosis & Bressler, 2007), which supports hypotheses suggesting the role of a supernatural agent to promote cooperation. Experimentally the picture is mixed (Hoffmann, 2013). Some studies find religiosity in general is positively associated with cooperation (Anderson & Mellor, 2009) while others find increased cooperation only between fellow members of the same religious group (Chuah et al., 2014). One field study found individuals who participated in extreme Hindu rituals made larger donations to a temple charity (Xygalatas et al., 2013). Overall, the evidence that religiosity per se increases cooperation is inconsistent, which casts some doubt on the supernatural punishment hypothesis.

Other work has instead emphasised the reputational benefits of displaying religiosity. Individuals who engage in displays of religious devotion could attain positive reputations and attract the most cooperative partners (Ge et al., 2019). In support of this theory, an experiment showed small communities (with more salient reputational effects) were more cooperative than larger communities and religiosity but not fear of divine punishment was associated with cooperation (Ge et al., 2019). This is consistent with ethnographic records of religious figures who are respected and revered rather than feared (Lenfesty & Morgan, 2019). As such, it is argued that religion is more likely to spread cooperation through a religious figure's prestige (section 2.3.2.4) rather than through dominance and fear of punishment (Lenfesty & Morgan, 2019).

2.2.3.5. Maladaptive by-product hypothesis

Some argue that human cooperation might be best explained as an evolutionary by-product or cognitive mistake (Morin, 2014; Schroeder et al., 2015). Using heuristics could lead individuals to cooperate when they otherwise should not (as predicted by game theory). This is said to be because of a mismatch between the ancestral conditions where humans evolved (repeated kin interactions) and modern cooperative contexts (Raihani & Bshary, 2015). Also, there may be greater costs associated with failing to cooperate in an iterated context than erroneous cooperation in a one-shot context (Delton et al., 2011). Relying on heuristics may represent a realistic approach to social dilemmas as they avoid the cognitively demanding (or sometimes, impossible) task of evaluating the costs and benefits of a given social interaction (Simon, 1990). This hypothesis is supported by a model, which showed that uncertainty in the length of prisoner's dilemma games favoured cooperation (Delton et al., 2011) and an experiment which found higher cooperation if participants were forced to make quick decisions (Rand et al., 2014).

Cooperation may have evolved because it is genetically linked to other traits. Selecting for one trait often results in changes in others, as was shown in the decades-long experiment domesticating wild foxes. Though the captive population was successfully bred to be tamer they also underwent a suite of other morphological and life history changes (Trut, 1999). It has been proposed that the genes coding for mutual tolerance in humans may have been genetically linked to those coding for generalised prosociality (Schroeder et al., 2015). This is similar to the self-domestication hypothesis, which suggests that humans evolved to be more prosocial by continually selecting prosocial partners. In support of this, authors have pointed towards similarities in human facial structure to other species that have undergone domestication (like dogs) (Theofanopoulou et al., 2017) and recent research has linked such facial morphing to the BAZ1B gene found in Williams syndrome, where individuals are characterised by hyper sociality (Zanella et al., 2019). Despite these alternative explanations, the scope and selectivity of human cooperation in many instances makes it hard to completely discount other accounts for cooperation such as reputation and reciprocity (Morin, 2014).

2.2.3.6. Proximate and demographic factors

In experiments, most individuals can be grouped into strong cooperators (often cooperate), free riders (rarely cooperate) and conditional cooperators (cooperate if others cooperate) (Kurzban & Houser, 2001). This alludes to individual differences between participants in their cooperative behaviour. Several other factors have also been correlated with cooperative behaviour.

One major consideration are demographic factors. The pattern associated with biological sex is complex. A large meta-analysis identified that overall cooperation rates do not differ between men and women, but there is some indication that males cooperate more with other males and females cooperate more than males in mixed sex interactions (Balliet, Li, et al., 2011). High scores on the personality traits agreeableness and conscientiousness (John & Srivastava, 1999) positively predicts cooperative behaviour (Kagel & McGee, 2014; Kurzban & Houser, 2001; Ross et al., 2003; Volk et al., 2011). At the country level, mean agreeableness predicted the number of sustainable policies enacted by the government (Hirsh, 2014). Honesty predicted cooperation in PD games where there was a strong temptation to defect, but not the SHG (Hilbig et al., 2018). Emotions like outrage and admiration for cooperative acts have also been shown to influence cooperation in a field context (van de Vyver & Abrams, 2015). Additionally, in economic games, priming guilt, but not shame, increased cooperation levels (de Hooge et al., 2007) whereas priming of anger decreased cooperation but increased punishment rates (Drouvelis & Grosskopf, 2016).

Another important consideration is variation between populations, as most experiments that theories were based on were conducted on WEIRD populations, usually students (Henrich, Heine, et al., 2010). Using various economic games, considerable variation has been documented across cultures in participant's levels of cooperation (Henrich, Boyd, & Bowles, 2005; Lamba & Mace, 2011). Later experiments showed this could partly be explained by a community's size, their degree of market integration and if they adopted a world religion (Henrich, Ensminger, et al., 2010). Rates of antisocial punishment also varies across societies. Societies with stronger norms of civic cooperation (positive attitudes towards taxation or not dodging travel fares) exhibited less antisocial punishment and higher levels of cooperation than societies that used more antisocial punishment and had weaker civic cooperation norms (Herrmann et al., 2008). Later analysis of the same data suggested the societies which used antisocial punishment were more likely to be ex-communist (Bruhin et al., 2020).

Variability also exists within societies. For example, cooperation rates across 16 populations of the same society varied substantially according to demographic factors such as village size and the presence of adult sisters within the village (Lamba & Mace, 2011). Within WEIRD populations, a trust game experiment found greater levels of trust, less use of punishment and greater consistency in behaviour in central executive officers than students (Fehr & List, 2004). Cooperation also varied between richer and poorer areas of a UK city, where the rich area exhibited higher cooperation and had less crime (Nettle et al., 2011). Group level variation is not specific to humans. After a tuberculosis outbreak killed many aggressive dominant males in a group of baboons, the group became less aggressive, and migrating individuals appeared to adopt this behaviour by observing surviving adult males (Sapolsky & Share, 2004). A recent field experiment identified variation between chimpanzee groups, such that more socially tolerant groups were also more prosocial which provides initial evidence that some animal prosocial behaviour may also vary between populations (van Leeuwen et al., 2021).

2.3. Social learning

The field of social learning has a similarly rich history to that of cooperation. From the understanding that indiscriminate copying is unlikely to be adaptive in most instances (Boyd & Richerson, 1995; Rogers, 1988), topics of study have shifted towards the consideration of social learning strategies or transmission biases. These strategies, biases or heuristics, influence when individuals use social learning, whom individuals learn from and what information is transmitted. Broadly, social learning strategies can be separated into context dependent strategies and content dependent strategies (figure 2.2). The functional definition of social learning I adopt in this thesis is the following: “Learning which is influenced by observation of, or interaction with, another [individual] or its products” (Heyes, 1994, p. 207). In the following section I briefly review the major social learning strategies described in the literature but grant the greatest attention towards the strategies explored in this thesis.

It should be noted that the evolution of social learning strategies does not require that individuals employ them consciously (Laland, 2004) and there is increasing appreciation that individuals do not necessarily apply strategies consistently or use only a single strategy at a time (Bono et al., 2018; Kendal et al., 2018). As such, studies which identify the use of any given learning strategy in particular scenarios do not also imply that no other strategies are ever employed in that context, particularly because of state based strategies such as those based on age or reproductive state. Furthermore, some researchers have argued of the danger of following the “behavioural gambit”, where psychological constraints are ignored in favour of evolutionary explanations for behaviour (Fawcett et al., 2013). This can complicate evaluations of the use of social learning strategies in some cases, as psychological constraints can often prevent animals from behaving optimally or selection may only favour solutions which perform well on average (Fawcett et al., 2013).

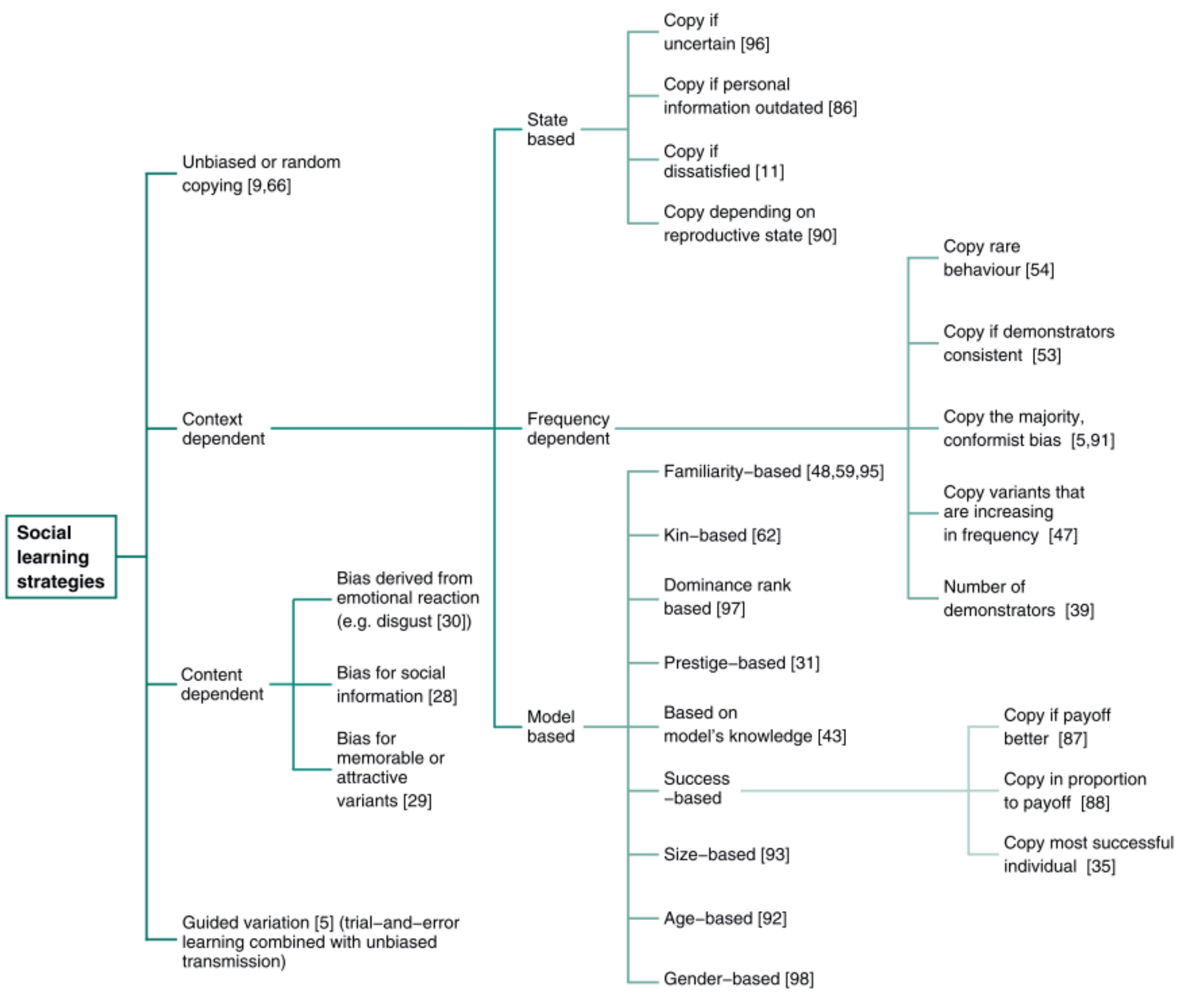


Figure 2.2*. Diagram showing various social learning strategies which have received theoretical support. Taken from* (Rendell et. al., 2011).

2.3.1. Content biases

Content biased transmission occurs through the selective attention of individuals towards certain kinds of information such that the frequency of that information persists or increases (McElreath & Henrich, 2003; Mesoudi, 2016). These biases may exist due to underlying cognitive mechanisms or be influenced by culture. There is evidence for both these possibilities. A colour naming experiment with pseudo words found participants' word groupings converged to those observed in other languages (J. Xu et al., 2013), indicating a cognitive disposition. Conversely, American participants were more likely to detect changes in focal objects in scenes while Japanese participants detected those in the background, indicating cultural differences (Nisbett & Miyamoto, 2005). Another common method to evaluate content biases are transmission chain experiments (Bartlett, 1932). In these experiments, participants are required to remember content then pass this along to the next individual in a chain, representing pseudo-generations (reviewed in Mesoudi & Whiten, 2008). This allows experimenters to identify what kind of information survives to the end of the chain. The proceeding sections (sections 2.3.1.1-3) cover three widely studied content biases, though this list is not intended to be comprehensive.

2.3.1.1. Socially relevant content

Proponents of the social brain hypothesis suggest a driver for the evolutionary pressure towards larger brains in humans was the cognitive demand associated with complex social interactions (Dunbar, 2009). This implies that human cognition should have evolved to attend to and transmit social information (Mesoudi et al., 2006). An early observational study identified a substantial tendency (as high as 70%) for conversations to discuss social information (Dunbar et al., 1997). Likewise, college students ranked receiving positive information about their friends or partners and negative information about rivals as the most appealing kinds of social information (Mcandrew et al., 2007).

Later experiments have supported a bias for socially relevant content. By the end of a transmission chain, information regarding social relationships ("gossip") was retained in greater quantity and more accurately than geographical information (Mesoudi et al., 2006). Similarly, urban legends containing social information were better remembered than control material but were not more likely to be chosen to transmit or receive (Tehrani et al., 2015). Recall of the social content of audio recordings was also better remembered than non-social content (Reysen et al., 2011). It has also been shown that adding normative statements (a socially relevant cue) to behavioural text made it easier for participants to recall the text (O'Gorman et al., 2008). This social bias may arise early in development, as indicated by a naturalistic study in a classroom, where gossip rather than general knowledge was much more frequently transmitted between children (McGuigan & Cubillo, 2013). Outside of a laboratory setting, our cognitive bias for social information can be exploited. Gossip or social content is a staple of newspaper headlines, who have historically paid huge sums of money for exclusive celebrity stories (McNamara, 2011). In a large analysis of 300 years' worth of front-page stories, stories surrounding altruism, cheating or violence have consistently been the most popular (Davis & McLeod, 2003). When asked to make newspaper stories more memorable, participants in the lab, among other things, added additional social content (Stubbersfield et al., 2018).

2.3.1.2. Emotionally relevant content

Like social content, emotional content has also been suggested to transmit more readily than other types of information. Cognitive neuroscience has long since established the key role of emotion in the processing and recall of episodic memory (LaBar & Cabeza, 2006; Tyng et al., 2017). Within the lab, many experiments have focused on exploring the type of emotional content that is transmitted more frequently.

Participants self-report they are more likely to pass on urban legends (C. Heath et al., 2001) and anecdotes (Peters et al., 2009) high in emotional content. Using transmission chains, greater willingness to transmit and receive and greater recall accuracy was found for urban legends containing highly disgusting content (Eriksson & Coultas, 2014). This general pattern of results also extended beyond disgust, to include amusement, interest, and surprise (Stubbersfield et al., 2017). Like social content, participants also manipulated the emotional valence of news stories to make them more memorable (Stubbersfield et al., 2018). In a field study, within 10 days of 33 students visiting a morgue, 881 individuals had learned of the visit and the number of people each student told was predicted by their self-reported emotional reaction (Harber & Cohen, 2005). There is also evidence that ambiguous information may be framed more negatively when it is repeated and negative information spreads more easily than positive information (Bebbington et al., 2017). Within applied settings, a review of New York Times articles found articles high in emotional arousal were more likely to be shared via email (Berger & Milkman, 2010) and false rumours of classical music enhancing children's intelligence spread faster in states anxious about the education system (Bangerter & Heath, 2004). Finally, Reddit users reported their voting decisions were directed towards posts high in empathy or humour (Priestley & Mesoudi, 2015).

2.3.1.3. Survival relevant content

Evolutionary psychologists have argued that there may exist a bias towards survival or fitness relevant information (Nairne & Pandeirada, 2008). Word recall tasks have shown that participants are better able to remember words that have some survival relevance (Kang et al., 2008; Nairne & Pandeirada, 2008). Students showed a greater ability to learn and recall unfamiliar animal tracks and cooking utensils than other categories of objects (Sharps et al., 2002). A different experiment showed participants were better able to remember survival relevant words when the context was ancestrally relevant (grassland environment) than a contemporary environment (city), which supported the innate cognitive bias hypothesis (Weinstein et al., 2008). Other evidence does not support a survival bias, however. Urban legends frequently contain social information, but rarely survival relevant information (reviewed in Tehrani et al., 2015) and an experiment showed urban legends with social content were more memorable than those with survival content (Tehrani et al., 2015). Therefore, it is unclear whether the memory advantage for survival information extends to transmittable cultural traits.

2.3.2. Context biases

Context biases refer to social learning strategies that bias transmission based on the characteristics of model(s) or the frequency of traits, rather than the kind of information (McElreath & Henrich, 2003). Strategies can be split into "when" strategies which influence when individuals use social learning as opposed to/in addition to asocial learning, and "who" strategies which influence from whom an individual will socially learn (Laland, 2004). Compared to content bias research, studies investigating context biases use a wider variety of methods including experiments and theoretical models. This thesis focuses specifically on "who" strategies, so these will be considered in greater depth than "when" strategies.

2.3.2.1. *"When" strategies*

The most straightforward when strategies are those which involve assessing the quality of, or risk associated with, one's prior information. If acquiring personal information is costly individuals tend to use social information. Thirty-six-month-old infants were more likely to adopt a demonstrated solution to a puzzle if they had struggled themselves or witnessed a different adult struggle with the task (Williamson & Meltzoff, 2011). Rats also used social information to inform their food choice if they were uncertain what food type had previously caused them to become sick (Galef et al., 2008). Foraging minnows adopted a socially demonstrated feeding patch when simulated predation was high versus low (Webster & Laland, 2008). Other experiments have shown that humans (Morgan et al., 2012) and primates (Kendal et al., 2009) will rely on social learning more heavily when the task is more difficult, when they are performing poorly (Atkisson et al., 2012) or when they are uncertain (Toelch et al., 2013). "When" social learning strategies have also been shown in insects, indicating they do not require complex cognitive mechanisms (Laland, 2004). Bean beetles go against their preference and choose occupied laying sites when offered only undesirable laying options (Otake & Dobata, 2018) and bumblebees adopt a demonstrator's flower preference after experiencing poor rewards from the other flower choice (Jones et al., 2015).

There is also evidence that social learning use varies based on phenotypic characteristics, either because some kinds of individuals are less certain of, or confident in, their own knowledge or less able to use asocial learning. Younger individuals have been shown to use social learning more and innovate less than older individuals and were less selective in their social learning use based on task difficulty or their own ability (Flynn et al., 2016; but see Leris & Reader, 2016). Similarly, high ranking chimpanzees changed their behavioural strategy less often than lower ranking individuals in response to social and asocial information, indicating less strategic information use (Kendal et al., 2015). In sticklebacks, reproductively active females (who are at greater risk from predation) relied more heavily on social information than males who were less affected by social information with respect to feeding preferences (Webster & Laland, 2011). Personality traits such as high extraversion and boldness also predict social information use (Mesoudi et al., 2016; Rawlings et al., 2017; Toelch et al., 2013). Social learning rates also vary by culture. Chinese participants relied on social information more heavily than UK participants in an artifact design task (Mesoudi et al., 2014) and the extent, and kind, of social learning that individuals used varied across 14 villages in India (Lamba, 2014).

2.3.2.2. Frequency dependent bias

Following a frequency dependent bias, individuals are inclined towards copying a trait based on its frequency within a population (Rendell et al., 2011). One such strategy is conformity[8] where individuals are disproportionately likely to adopt the most common trait in accordance with the assumption that successful traits will be present in the greatest frequency (Henrich & Boyd, 1998; Whiten, 2019a). Many theoretical models have supported the efficacy of conformity for adopting beneficial traits (Hastie & Kameda, 2005; Henrich & Boyd, 1998). Others have argued that strong conformity may inhibit the spread of innovated traits, as conformity would ensure these traits were never copied (Kandler & Laland, 2009; Whitehead & Richerson, 2009).

Two other further qualifications are important here. Conformity in the cultural evolution sense (which results in common traits increasing in frequency) follows a characteristic “S” shape where common traits are overrepresented (Claidiere & Whiten, 2012) rather than copied at a rate equivalent to their representation in the population (random copying). Secondly, a distinction is often made between copying the most common behaviour in the population versus the behaviour expressed by the most individuals (Morgan et al., 2019; van Leeuwen et al., 2016). In an experiment, children and chimpanzees posted tokens in a slot where three different individuals had posted a single token rather

[8] Some have also considered a non-conformity or anti-conformity, where individuals disproportionately copy the least common traits.

than the slot where one individual had posted three tokens (Haun et al., 2012). This distinction can be important, as the kind of conformity employed can have different population level consequences in terms of the fixated trait and how successfully populations adapt to changing environments (Morgan et al., 2019). A disproportionate tendency to copying individuals (rather than behaviours) in the majority represents following collective intelligence.

Experimentally, the evidence for conformity is mixed. There are several experiments which have shown human participants use frequency information in a variety of different tasks (Efferson et al., 2008; Morgan et al., 2012; Muthukrishna et al., 2016) but others that find little influence of frequency information on participants choices (Eriksson & Coultas, 2009) and that other types of social information are often preferred (Barrett et al., 2017; McElreath et al., 2008; Mesoudi, 2011). The use of conformity also varies alongside contextual factors, being more common when individuals are faced with more choices (Muthukrishna et al., 2016) or are more uncertain (Morgan & Laland, 2012). However, of the experiments that find use of frequency information, relatively few demonstrate disproportionate copying of common traits (but see Morgan et al., 2015). Studies investigating conformity are also usually restricted to binary traits (but see Morgan & Thompson, 2020).

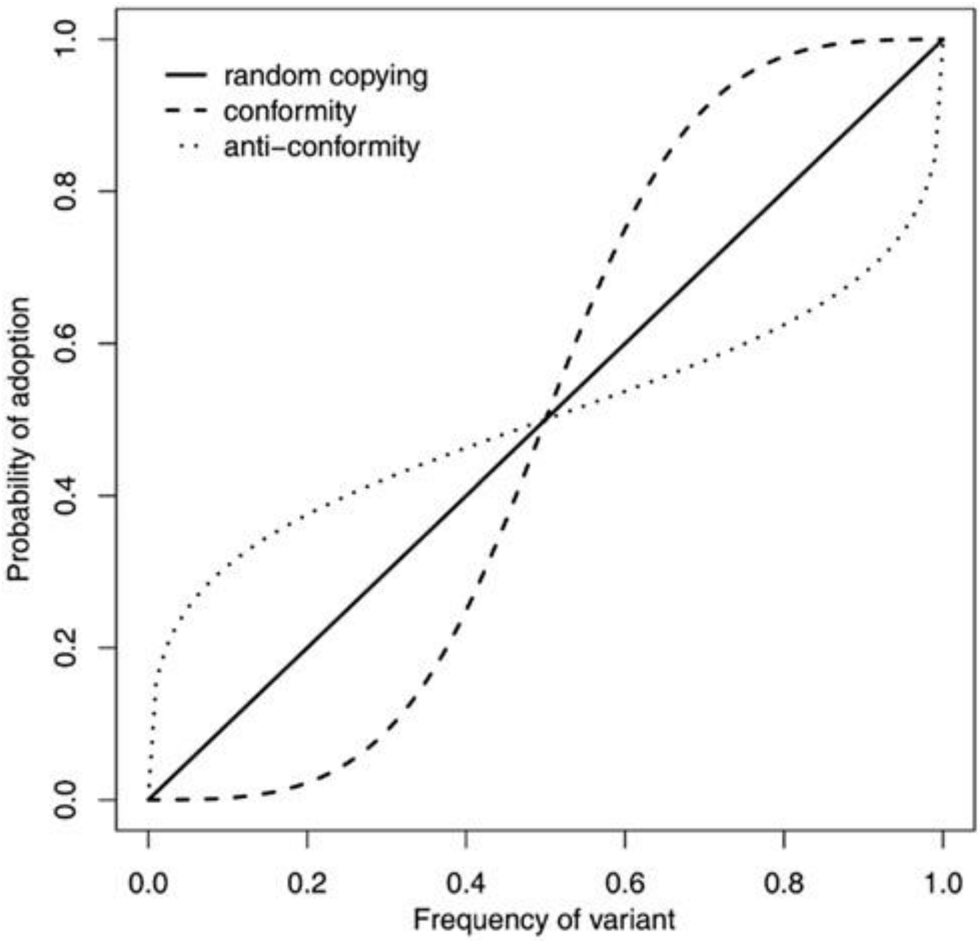


Figure 2.3. *Diagram contrasting random copying (solid line), conformity (dashed line) and anti-conformity (dotted line) (taken from Morgan & Laland, 2012).*

2.3.2.3. Payoff bias

Using a payoff bias, individuals are inclined towards copying a behaviour, displayed by an individual, that has a greater payoff than themselves[9] (McElreath & Henrich, 2003). The spreading of traits according to their fitness (or payoffs) is usually a default assumption in evolutionary models, termed the replicator dynamic (Hofbauer & Sigmund, 2003; Taylor & Jonker, 1978). This is advantageous, as avoiding heuristic strategies like conformity or prestige bias (section 2.4.4.) reduces the chance of copying outdated or maladaptive traits. Formal modelling has shown that if individuals copy

[9] It is somewhat unclear where to place a payoff bias. While often it is thought of as a context (who) bias as copying is directed towards individuals who are doing better it can also be thought of as a content (what) bias or a when bias (Rendell et al., 2011). Whichever definition is chosen, the functional predictions remain the same.

strategies weighted by their payoffs, this can outcompete other strategies (Schlag, 1998). The adoption of continuously improving cultural traits is also a requirement for cumulative cultural evolution (Mesoudi & Thornton, 2018).

A payoff bias can operate in different ways, which in turn can affect its overall success at spreading beneficial traits. For instance, a learner can copy the trait with the best average payoff or the trait with the best single payoff (Baldini, 2013). A model showed that both strategies are susceptible to the fixation of suboptimal traits depending on their frequency and the variance of population trait payoffs (Baldini, 2013). Consequently, they can be outcompeted by unbiased copying. An alternative possibility is that learners sample a small group of individuals from the population (rather than consider the whole population) and adopt the best performing trait from this group (Takahashi & Ihara, 2019). This too is vulnerable to inferior traits and is unstable against unbiased copying without some amount of asocial learning (Takahashi & Ihara, 2019). The success of payoff based social learning may depend on environmental conditions and may be suboptimal when the payoffs associated with traits are highly variable (Baldini, 2013; Takahashi & Ihara, 2019).

Empirically, there is evidence of the use of payoff bias in several animal species. Sticklebacks switched to an observed food patch only if it provided a greater return (Kendal et al., 2009). In a token exchange task, children appeared to use a form of 'copy if better' strategy involving comparing personal payoff and social payoff information, but chimpanzees showed little use of social payoff information (Vale, Flynn, et al., 2017). In wild capuchins, methods to extract a novel food item spread through payoff and age biased copying but not conformity (Barrett et al., 2017). Conversely in humans, despite its superior performance, a payoff bias was rarely used in an arrow designing task (Mesoudi, 2011). Though in field contexts, payoff biases have been shown to predict opening moves in the game of GO (Beheim et al., 2014) and managers formation choices in football (Mesoudi, 2020).

2.3.2.4. Prestige bias

Employing a payoff bias requires that individuals can accurately assess the relative payoffs of different behaviours displayed by models (McElreath & Henrich, 2003). Often this is difficult, either due to cognitive limitations or inadequate access to this information. One solution is to rely on a heuristic strategy and copy generally successful individual. Prestige is one such cue (Henrich & Gil-White, 2001). Individuals who display generally useful skills are conferred prestige through voluntary deference by followers who do so to gain access to this information. Uninformed observers can then use primary (success in other domains) or secondary (popularity) cues to decide whom to copy, on the assumption that this information is better than average (Jiménez & Mesoudi, 2019). The two key features of prestige are its contrast to dominance (as prestige is conferred rather than asserted) and that prestigious individuals are influential beyond their domain of expertise (Henrich & Gil-White, 2001). This is regularly exploited by advertisements using celebrity endorsement (Mukherjee, 2009) but it can result in the spread of maladaptive behaviour (Mesoudi, 2009). Prestige should also be used primarily when payoffs are difficult or impossible to assess (Brand et al., 2020; McElreath & Henrich, 2003).

Most of the evidence for prestige biases comes from humans (reviewed in Jiménez & Mesoudi, 2019). For example, residents in a Fijin village learned about medicinal plants from individuals knowledgeable in yam growing as payoffs related to medicinal plant use are ‘fuzzy’ (Henrich & Broesch, 2011). In the royal navy, two studies showed that team performance and the flow of information was correlated with the degree of prestige conferred upon leaders (Offord et al., 2016, 2019) and participants in an arrow designing task copied the individuals that other participants had spent more time looking at, a cue of deference/prestige (Atkisson et al., 2012). Children also showed a greater tendency to imitate causally irrelevant actions of a high-status than a lower status individual (McGuigan & Cubillo, 2013; but see Chudek et al., 2016). A series of recent experiments supported the prediction that prestige cues are used only when payoff information is unavailable (Brand et al., 2020), but found little evidence that participants copied individuals outside of the prestigious individual's domain of experience (Brand et al., 2021). This lack of domain general prestige biased copying was also shown in children (Chudek et al., 2012). Prestige was not found to be influential in cases where no expertise was required for judgements, for example in the rating of the likeability of quotes (Acerbi & Tehrani, 2018) or the judgement of art (unless the participant was an art expert) (Verpooten & Dewitte, 2017). The evaluation of an individual's prestige may also depend on the observer. Political leaders are more likely to be perceived prestigious if their political beliefs align with the observer (Jiménez et al., 2021). Like payoff learning, this suggests that the influence of prestige may depend on contextual factors and whether participants perceive the model as prestigious.

2.4. Cooperation and social learning

There is a small, but growing, field of research focused on investigating the extent that social learning affects cooperative behaviour. Before reviewing this research, it is necessary to introduce a final account for human cooperation that explains how social learning can promote cooperative behaviour. The theory of cultural group selection (or sometimes, multilevel selection) suggests that intergroup competition might provide a mechanism for the maintenance of cooperation[10]. If multiple groups in a population compete for resources and cooperation predicts success in this competition, cooperation could spread through the population through social learning or assimilation (Boyd & Richerson, 2009; Henrich, 2004; Richerson et al., 2016; Smith, 2020). Cultural group selection requires other mechanisms which stabilise group behaviour, be it social learning, punishment or social norms. Once group level variation is stabilised, between group variation may then be sufficiently strong to allow group, rather than individual, beneficial behaviour to evolve (D. Smith, 2020). Advocates of this approach have argued that it provides the most compelling explanation for human cooperation

[10] Note that multilevel or cultural group selection is different to classic group selection (Wynne-Edwards, 1963). Classic group selection fails because a single selfish individual can invade a population of altruists. Cultural group selection avoids this by suggesting that individual and group success is correlated, and that cooperation is a cultural trait which is relatively resistant to migration.

(Richerson et al., 2016) as it does not require that individuals are related and can explain the initial emergence of cooperation (Boyd & Richerson, 2009b).

Cultural group selection is, however, not without criticisms. For one, it requires that group conflict favours cooperation. Although this assumption has been supported by laboratory studies (Majolo & Maréchal, 2017; Puurtinen & Mappes, 2009a), there is less evidence for this in field settings (Mace & Silva, 2016; Silva & Mace, 2014). Other critics point out that because cultural group selection requires that cooperation is a selectable group level trait, to reject individual level explanations for cooperative behaviour, links must be demonstrated between group and individual fitness (Barclay & Krupp, 2016; Krasnow & Delton, 2016). Cooperation is also highly variable, even within supposedly culturally homogenous groups (Nettle et al., 2011) shedding doubt on the extent to which cooperation could be a selectable group level trait. Principally however and of specific interest for this thesis, it is not yet clear the extent to which social learning can spread cooperative behaviour between individuals (D. Smith, 2020).

In the proceeding section and the remainder of this thesis, I will evaluate and explore the extent to which social learning influences behaviour in social dilemmas. Though direct testing of any of the assumptions of cultural group selection is beyond the scope of this thesis, the findings from this thesis may provide further insight into the discussion. I focus on three social learning strategies: payoff bias, conformity and prestige bias. These strategies are chosen as they are the most widely studied social learning strategies within the social learning literature and have already (partially) been considered alongside cooperation (Burton-Chellew, el Mouden, et al., 2017; Burton-Chellew & Amico, 2021; P. van den Berg, Molleman, & Weissing, 2015). However, little work has addressed the role of prestige bias within cooperative scenarios and no work has considered all three strategies simultaneously. In the proceeding sections, though I focus primarily on these three strategies, I also draw upon research which has considered social influence more generally or that which considers multiple strategies simultaneously.

2.4.1. General social influence

Many studies have investigated social learning without focusing on any specific strategy or have considered the influence of other individuals in a very broad sense. A selection of this research is discussed below.

Theoretical results concerning broad social learning are somewhat mixed. One model found that social learning could sustain cooperation within structured populations (Quan et al., 2018) but others find that this is reduced in uncertain environments (Andras, 2018) or that if cooperation did not affect reproductive success then social learning favoured spite (Lehmann et al., 2008). This also varied by game type, as social learning was favourable in a stag-hunt game, but asocial learning was more favourable in the snowdrift game (Danku et al., 2018).

Understanding of the effect of broad social influence on cooperation comes primarily from field studies. A classic experiment found that helping to fix a flat tyre and donations to a charity box were more common when the public had witnessed another individual engaging in the behaviour (Bryan & Test, 1967). At Zurich University, students were more likely to continue contributing to a student fund if they believed others were doing the same (Frey & Meier, 2004a). In the Hadza tribe, donations of honey were predicted from donations offered by other individuals and there was greater variation in donations between camps than within camps (as cooperators tended to assort with fellow cooperators) providing indirect evidence of social learning between social groups (Apicella et al., 2012). Cleaner fish were also shown to use social learning to both avoid, or behave more cooperatively towards, feeding sites (simulating client fish) that punished and to cheat more when they would not be punished (Truskanov et al., 2020). In a laboratory public goods game (PGG) with repeated one-shot interactions, participants continued to cooperate after interacting with a cooperator causing cooperation to spread through the social network (Fowler & Christakis, 2010). Participants also generally cooperated more when placed in groups with known cooperative individuals (Gächter & Thöni, 2005).

Individuals may also be influenced by social norms. Norms are often separated[11] into descriptive (what do most people do?) and injunctive (what do most people approve of?) norms (Cialdini et al., 1991). Supporting the distinction, both kinds of norms often influence individuals differently (Wechsler et al., 2003). Humans readily follow social norms and often invoke them automatically in contexts where they are established (Aarts & Dijksterhuis, 2003) and enforce them on others from an early age (Schmidt & Tomasello, 2012). However, individuals often underestimate the extent that social norms influence their behaviour (Frey & Meier, 2004b; Nolan et al., 2008). Within cooperation, social norms are agued to be a key factor for cultural group selection (Richerson et al., 2016) and the resolution of social dilemmas more broadly (Bicchieri, 2005; Ostrom, 2000).

Descriptive social norms function similarly to conformity (see section 2.4.3.), so this section will primarily focus on injunctive social norms, though many studies consider both simultaneously. Several large-scale sustainability studies have found that while injunctive norms are often effective to reduce energy consumption (Schultz et al., 2007) and increase recycling (Thøgersen, 2008), descriptive norms can cause a boomerang effect, where individuals converge towards the average behaviour (but see Allcott, 2011). This is supported by a US survey and experiment which found second order norms (what does the country think) more strongly predicted saving energy than a participant's personal norms (Jachimowicz et al., 2018). Guests were also more likely to reuse towels in response to normative information than environmental information (Goldstein et al., 2008) and drivers to switch off their engines when queueing when the normative prompt was relevant to their social group than when it was not relevant (Player et al., 2018). However, social norms are not always effective. Injunctive prompts did not increase the chance that individuals would donate the money they received from returning plastic bottles to a pro-environmental charity (Bergquist et al., 2020; Neumann, 2019). One explanation is that the expected normative behaviour was already clearly understood (Bergquist et al., 2020), suggesting that normative prompts may be more effective when

[11] A further classification sometimes used is between moral (norms which prevent harm) and conventional (norms which are purely tradition) (Eggleston & Turiel, 1983). This distinction is not considered in this review.

the expected norms are not clear. Social norms are also often less impactful overall than social influence interventions that employ block leaders (specific individuals who volunteer to disseminate information Abrahamse & Steg, 2013).

2.4.2. Payoff bias

Very little work has focused exclusively on a payoff bias in social dilemmas. This is likely because the transmission of behaviours with higher fitness is a standard assumption of evolutionary models (Taylor & Jonker, 1978). Given that defectors invariably achieve higher payoffs than co-operators within PD games, it follows that a payoff bias would result in the decline of cooperation. This generally plays out, as in the absence of other mechanisms, defection emerges as the ESS in PD games (Doebeli & Hauert, 2005a). If a payoff bias is considered, it is usually in contrast to other social learning strategies or mechanisms.

Experiments can still provide valuable insights into the role of a payoff bias in cooperation as individuals do not always behave consistently with theoretical models. In two closely related experiments, individuals played iterated PGGs and were provided feedback about how their behaviour was affecting the payoffs of the other participants, they responded by reducing their contributions (Burton-Chellew et al., 2015; Burton-Chellew & West, 2013). A recent study comprising of an experiment, model and meta review of 237 PGG studies found declines in cooperation was predicted by the degree of influence participants had on their own payoffs, which is consistent with the use of a payoff bias (Burton-Chellew & West, 2021). Conversely, a payoff bias can occasionally benefit cooperation. In an experiment with free movement between groups, participants used payoff learning to migrate into cooperative groups with punishment (Gurek et al., 2006). This shows that, although a payoff bias will erode cooperation acting in isolation, it may not be detrimental for the evolution of cooperation alongside other mechanisms.

2.4.3. Conformity

Conformity could benefit the evolution of cooperation in two ways. Provided cooperation is sufficiently common, conformity could act directly on cooperative behaviour resulting in an increase in its frequency and allow it to resist invasion from defectors (Boyd & Richerson, 1985). Conformity could also act on social norms within a group, which could themselves promote a norm of generalised reciprocity[12], or homogenise other unrelated norms which may make the spread of cooperation more likely (Molleman, Pen, et al., 2013). Most research has focused on the direct copying of cooperative behaviour (but see Gächter & Thöni, 2004).

Theoretical models typically find that the benefits of conformity depend on its strength and the composition of the population. In PD games, conformity speeds up the growth of defectors in the population, if they are initially common (Carpenter, 2004) and undermined the evolution of cooperation when it was initially rare (Lehmann & Feldman, 2008). But it can help cooperation to evolve if the temptation to defect is low (Cui & Wu, 2013) or if it coevolves with punishment (Andrés Guzmán et al., 2007). In the SD game, which usually favours stable populations of defectors and cooperators, including conformity resulted in pure cooperation or defection, whereas in the SHG, conformity was not used (Molleman, Pen, et al., 2013). Overall, it has been argued that the influence of conformity in the evolution of cooperation is likely to be small (Molleman, Pen, et al., 2013).

Empirical patterns differ from theoretical predictions in that many studies find at least some evidence for a beneficial effect of conformity, despite the effects being small. The caveat being that these studies do not provide evidence of cultural evolution conformity where traits are disproportionately likely to be copied (Henrich & Boyd, 1998). This means that it is not possible to distinguish whether

[12] Generalised reciprocity is a mechanism where individuals cooperate based on the behaviour of their previous partner. Generalised reciprocity is cognitively easier to apply (not requiring partner recognition) and has been suggested as a potential mechanism for small groups (Pfeiffer et al., 2005).

participants are employing unbiased copying (which also predicts that common traits are copied more frequently) or conformity. In each case below, the way conformity was operationalised is described.

Two experiments using a PD game investigated the influence of conformity compared to reciprocity. In the first experiment, participants received information about the donations of their own group which affected their payoff (reciprocity) or also an external group which did not (conformity) before making a one-shot decision. Participants were more cooperative in response to high donations than low donations[13] from both groups, but the effect associated with reciprocity was larger than conformity (Bardsley & Sausgruber, 2005). In the second experiment, before choosing how many of 100 tickets to donate to a partner in a two player PD, participants received information about their partner's previous behaviour towards the rest of the group (reciprocity) and their groups behaviour towards other individuals (conformity). Individuals donated more tickets if they received information suggesting their group/partner had been cooperative (donating most or all of their tickets) than uncooperative (donating very few tickets) but reciprocity was more influential (Romano & Balliet, 2017). On online fund-raising pages, individuals use previous donation amounts to determine their donations and a £10 increase in the mean donation increases subsequent donations by £2.50 (Smith et al., 2015). Adding a normative prompt along with tax reminders ("*9 out of 10 people pay their tax on time*") was also effective at raising tax compliance (Onu & Oats, 2014).

Other studies have found that conformity can increase free riding. Contributions in PGG experiments decline faster when participants have access to their group mates' behaviour as participants appear to desire to contribute less than their group (Neugebauer et al., 2009). Young children would also conform to a majority who chose not to donate a cereal bar, unless the recipient was described as hungry (and therefore more needy Engelmann et al., 2016). In a dice throwing experiment where participants self-reported their results, participants were more likely to lie and report higher rolls if

[13] The experiment compares the average donation of participants in response to "low" and "high" donations from the groups. It does not detail what a "low" and "high" donation is taken to be (Bardsley & Sausgruber, 2005).

they had seen social information suggesting that other participants had also lied about their dice rolls (Diekmann et al., 2015). Other field studies have demonstrated the “broken window effect” where overt signs of norm breaking predicts further norm breaking. For example, individuals were more likely to litter if the environment was covered in graffiti and ignore a no trespassing sign if other individuals had ignored requests to not park bikes to a gate (Keizer et al., 2008). Such findings point towards a more mixed picture than theoretical models with respect to the influence of conformity on cooperative behaviour. However, it remains that none of these studies provide evidence of the disproportionate copying that is characteristic of conformity (Henrich & Boyd, 1998).

2.4.4. Prestige bias and leadership

Leadership (or prestige) is potentially beneficial for the evolution of cooperation as it may be able to promote and spread cooperation, regardless of the initial composition of the population. Additionally. leaders are incentivised to cooperate and use responsible punishment to maintain their prestige (Henrich & Gil-White, 2001). Some have further speculated that leadership may have specifically evolved to solve coordination problems between individuals (King et al., 2009; Vincent, 2017). Providing support for the prestige hypothesis, although dominance predicted an individual becoming a leader, motivation and knowledge were also strong predictors (King et al., 2009). Many studies have investigated the role one-to-many transmission may play in cooperative dilemmas, though research is lacking compared with other social learning strategies.

Theoretical models have generally pointed towards prestige (or leadership) being effective to promote cooperation. In a large, repeated PD model which permitted agents to change their cooperative partner if they were dissatisfied, the population either reached a stable state of cooperation or defection but cooperative populations could be sustained by cooperative leaders, provided their payoffs were higher than defective leaders (Zimmermann & Eguíluz, 2005). A different model investigated a context where a single individual was copied with a varying probability (indicating their prestige) by their group mates. A larger prestige bias within a group was associated with larger groups where cooperation could be sustained which, in turn, favoured the leader to continue cooperating (Henrich et al., 2015). Compared to conformity, prestige could sustain cooperation to much greater levels, even under migration or when the initial population contained only few co-operators (Lehmann & Feldman, 2008).

Many experiments investigating leadership do so by selecting an individual at random to decide on their contribution first. In a PD game where a single individual (the leader) makes a cooperative decision before other group members (the followers), high donations from leaders are copied by followers and followers that are more cooperative continue to behave cooperatively when assigned to the leader role (Gächter et al., 2012). Groups with a prosocial leader (identified by questionnaire) were more cooperative and the leader used less punishment than if the leader was selfish or in groups without a leader (Harrell & Simpson, 2016). Ethnographic evidence also points towards fairness, charisma and generosity in group leaders in small scale societies as positive predictors of group performance, as leaders punished free riders without exploiting the spoils for themselves (von Rueden et al., 2014). An experiment supported this, finding that describing a leader as fair and charismatic through vignettes increased participants contributions to a public good (De Cremer & Van Knippenberg, 2002). Within the business world, successful leaders tend to be charismatic, fair and supportive, which is associated with increased group performance and productivity (Muteswa, 2016). Though, leadership is not associated with cooperation in every case. In the weak link game (lowest contribution determines payoff), leaders made only small contributions which subsequently elicited small and inconsistent increases in their group's cooperation (Cartwright et al., 2013) and a water management framed PGG found no evidence that participants copied leaders or that leaders were more cooperative than other individuals in the group (O'Garra & Alfredo, 2019).

Fewer experimental studies have focused specifically on the role of prestige or the traits associated with prestige, which is surprising given the predictions of theoretical models (Henrich et al., 2015). Participants were more cooperative in a dictator game after watching TED talks of charismatic speakers (Grabo & van Vugt, 2016) and some evidence was found that reading vignettes of selfish politicians made participants less able to understand ethical considerations of fictional scenarios (Watts et al., 2018). In the one experiment which evaluated status in a cooperative game, participants completed a general knowledge quiz which rewarded high scorers with visible stars before playing a social dilemma game. Starred participants were more frequently copied and made generally higher contributions than non-starred participants (Kumru & Vesterlund, 2010). Status shows mixed results within field contexts also. In the royal navy, team performance (requiring a degree of cooperation) and the flow of information was correlated with the degree of informal prestige conferred to leaders (Offord et al., 2016, 2019) but, despite common marketing strategies, many real-life cooperative behaviours are not influenced by celebrity endorsement (John et al., 2019; Wood & Herbst, 2007). These studies provide initial support towards theoretical models which suggest that prestige or status can be beneficial for the spread of cooperative behaviour, but research within this realm is lacking.

2.4.5. Multiple strategies

Although social learning strategies are often investigated in isolation, there is good evidence that individuals use them synergistically or interchangeably (Kendal et al., 2018). Studies investigating multiple social learning strategies are useful because they permit investigations of the possible interactions between strategies and whether there are preferred sources of information in social dilemmas.

Many studies consider payoff learning alongside other social learning strategies to explore the extent to which the latter can resist the deleterious effect a payoff bias has on cooperation. In a spatial PD game, moderate levels of conformity sometimes allowed defectors to be displaced by cooperators, despite payoff biased copying between neighbours (Jiang et al., 2015). A different model showed payoff bias eroded cooperation in the SD game when the temptation to defect was high but prevented the invasion of defectors in a SHG after cooperation was established through conformity (Szolnoki & Chen, 2018). A large multi group model considered differences between selection by contagion (slow copying of successful groups) and replacement (takeover by successful groups) where individuals could also use a conformity, payoff or prestige bias to update their behaviour. Under replacement, prestige and conformity could stabilise cooperation but under contagion, only a prestige bias could sustain cooperation as conformity and payoff bias ensured that rare co-operators could never invade a population of defectors (Molleman, Quiñones, et al., 2013). This supports previous evidence showing prestige was more effective in larger groups than conformity (Lehmann & Feldman, 2008). These results suggest that, of conformity, prestige and payoff bias, prestige bias represents the most likely strategy which could promote within group cooperation under cultural group selection (Richerson et al., 2016).

Amongst the only experimental work to compare multiple social learning strategies within an economic game context was conducted by Burton-Chellew and colleagues. One experiment showed little change in cooperative behaviour in response to frequency information but sharp declines in response to payoff information (Burton-Chellew, El Mouden, et al., 2017). Two earlier studies offered participants the choice of viewing information about the behaviour or the behaviour + payoffs of other participants in 4 different games. Though consistent variation was shown between participants in their preferred source of information, of those who opted for payoff information, individuals frequently switched their behaviour in response to it (Molleman et al., 2014; P. van den Berg, Molleman, & Weissing, 2015). This was also consistent with a recent experiment which found participants decreased their cooperation to a greater extent in response to success weighted information than common behaviour and a preferred to view success weighted information when offered the choice (Burton-Chellew & Amico, 2021). However, a cross cultural study found that the extent that individuals used conformity over a payoff bias depended on a variety of contextual factors including village size and the size of their social network (Lamba, 2014). This supports findings within cultural evolution more generally which document individual and cultural differences in social learning use (Kendal et al., 2018). Overall, payoff weighted information results in the decline of cooperative behaviour in social dilemmas and is preferred to other kinds of social information.

2.5. Application of cooperation & cultural evolution to sustainability

Sustainability is a fundamentally important topic at the individual and society level within the 21st century. Although the UK public's concern for the environment continues to grow (M. Smith, 2019) there remains a gap between positive environmental attitudes and sustainable behaviour (White et al., 2019). Sustainability, and the impact of climate change, can be understood through the lens of social dilemmas. The problems created by greenhouse gas emissions are spread across the whole world and solutions to such problems require collaborations between nations (Keohane & Victor, 2016). This is consistent with the PD because the benefits of the public good (the environment) are enjoyed by all parties, regardless of their contribution (reducing emissions) (Keohane & Victor, 2016). Climate change is particularly challenging because costly mitigation efforts benefit future generations and there is no world government to punish nations that do not reduce their emissions (Carattini et al., 2019). Other environmental issues can be interpreted differently. For example, Ozone depletion was more like a SD game, because CFC emissions disproportionately affected the nations that emitted them, meaning there was a stronger individual incentive (from the nations perspective) to cut their own emissions (Zefferman, 2018).

A growing field of research has been exploring ways that environmental sustainability can both be understood, and promoted, by applying insights from cultural evolution and the evolution of cooperation (Brooks et al., 2018). Cultural group selection offers a plausible mechanism by which sustainable behaviour could spread and be maintained within groups (Brooks et al., 2018). However, guidance on how research should approach the study of cultural group selection in applied settings has been a relatively recent effort (Kline et al., 2018)

Institutions to manage sustainability have been argued to evolve spontaneously as a product of cultural evolution (Ostrom, 1990) though it is not clear where these institutions come from (Waring et al., 2017). Modelling showed that cultural group selection promoted the evolution of institutions of private property, collaborative production, and conservation norms, following an initial population crash (Waring et al., 2017). Other research has focused on case studies where cultural group selection has appeared to promote sustainable behaviour. In California, wine production results in negative environmental consequences, but growers are often part of institutions that promote sustainable practices (Hillis et al., 2018). Using interviews and questionnaires, a study found evidence that the environmentally sustainable practices of successful growers were copied and failing businesses merged with or migrated into more successful businesses, which are features of cultural group selection (Hillis et al., 2018). In Maine, territory-based institutions in the form of harbour gangs formed around lobster farming, which were then defended using retaliatory aggression such as cutting rival traps (Waring & Acheson, 2018). For conservation, there was some evidence that sustainable practices (such as escape vents for undersized lobsters) spread through social learning or were imposed within institution boundaries (trap limits), but others such as size restrictions for saleable lobsters were externally imposed (Waring & Acheson, 2018). Also in Maine, Hupper et al., (2019) suggested policy changes based on cooperation research to encourage local producers, which included more generous loans and promoting worker cooperatives.

On much smaller scales, many studies investigating sustainable behaviour employ strategies (intentionally or otherwise) that are effective in traditional cooperation contexts. Residents of an apartment block were more likely to agree to reduce their use of air conditioning when they made public rather than private commitments, which follows predictions from indirect reciprocity (Yoeli et al., 2013). Similarly, in a game where participants donated to a climate fund, donations were higher when participant made donations publicly and those that contributed to the climate fund were rewarded by other participants in a different game (Milinski et al., 2006). Interventions may aim to make sustainability more convenient (Carlson, 2001), offer financial incentives for behaviour (L. Xu et al., 2018) or subsidies for housing refits (Durham et al., 1988), which eases the costs (time or monetary) associated with sustainability such that it no longer represents a social dilemma. Social norms are also widely studied in relation to sustainability (discussed in section 2.4.1) and have been argued to represent a powerful untapped potential to promote sustainability, particularly if this makes sustainability more visible (Carattini et al., 2019).

What is less clear, is the extent that individuals use social learning strategies within a sustainability context. There is some evidence that one-to-many transmission does facilitate the spread of sustainable practices. For example, the institutions which formed around California wine growers promoted many of their sustainable practices through one-to-many transmission (Hillis et al., 2018). In Sweden, the success of a wetland conservation programme was driven largely by a single influential individual (Olsson et al., 2004). A model showed that sustainable behaviour spread more easily though one-to-many transmission than through conformity, which could stabilise sustainable or unsustainable behaviour, depending on its initial frequency (Buenstorf & Cordes, 2008). Other research has shown that the promotion of sustainable clothing rental websites by Instagram "micro celebrities" increased the uptake in their followers (Shrivastava et al., 2020) and familiarity with Greta Thunberg (a well-known climate activist) predicted intentions to take action to reduce global warming in a representative US based sample (Sabherwal et al., 2021). To my knowledge, no studies have investigated social learning strategy use within a sustainability context. This represents an interesting theoretical gap and is the focus of chapter 6.

2.6. Conclusion

Cooperation and social learning represent fundamental traits for the generalised success of humans throughout our evolutionary history. As such, they have garnered extensive research attention across a variety of disciplines. In this review I have highlighted this breadth of research and shown that cooperative behaviour can be influenced by myriad of (likely, interacting) factors. Social learning is commonly used by human and non-human animals to inform behaviour in a variety of circumstances. Both cooperation and social learning are also influenced by individual and cultural variation. Research applying social learning to cooperative dilemmas, although new, has identified some general trends. Overall, there is strong support for the use of a payoff bias but the evidence surrounding conformity is mixed. Individuals readily use payoff information and reduce their cooperative behaviour, but often ignore conformity information. In models, conformity may increase or decrease cooperation depending on the initial composition of the population. Bold predictions regarding prestige are made from multiple theoretical models yet lack good empirical support.

There remain several interesting theoretical gaps within this field which I aim to address and build upon throughout this thesis. In the first empirical chapter (chapter 4), I expand upon the approach taken by experiments, described above, and experimentally assess the extent to which individuals use conformity, payoff bias and prestige biase to influence their cooperative behaviour within PD and SD games. This experiment represents the first of its kind to permit the free use of all three of these strategies, between PD and SD games. The experiment is also supported by a theoretical model, to investigate the consequences of my experimental results across a longer timeframe.

In the second empirical chapter (chapter 5) I focus on spite, a frequently neglected aspect of social behaviour. Little is known about the factors which influence spiteful behaviour and the possibility that spite may be influenced by social learning has never been directly addressed. This study represents a novel assessment of the extent that individuals use social learning in their decision to enact spiteful behaviour.

Chapter 6 steps beyond experimental settings and investigates social learning strategies (and social influence more broadly) in an applied cooperative setting (environmental sustainability). Within sustainable behaviour, the role that others may play has been investigated previously, usually by way of social norms and applied interventions. However, little direct attention has been given to specific social learning strategies, nor has any experiment investigated them in real life cooperative behaviour. This study applies a mixed methods approach, combining quantitative and qualitative methods to investigate the factors that affect sustainable behaviour with specific attention towards the role of conformity, prestige bias, and payoff bias. To this end, I employ a questionnaire with follow up semi-structured interviews and a behavioural measure with experimentally manipulated vignettes.

Chapter 3: General methods

In this chapter, I will provide an outline of the types of methods used within cooperation research. Because of its breadth, several different methodological approaches have been used within the literature, each offering unique advantages. I will therefore briefly discuss some of these approaches and describe those that are used in this thesis and offer some justifications for their inclusion. The section begins with discussions on laboratory experiments and how (typically) economic games have been used to investigate cooperative and spiteful behaviour. I will then offer some potential limitations of this approach and discuss how field studies and theoretical modelling might help to overcome these weaknesses. I will then describe the precise methods used in this thesis along with a discussion of the overall statistical approach

3.1. Approaches to studying cooperation

Over the last several decades, our understanding of the ways and mechanisms by which cooperation can evolve and the cases where individuals cooperate has steadily developed. To this end, a wide variety of methods have been used to address a range of different questions. These include experimentation (both laboratory and field-based) and theoretical models, which have been applied in a countless number of studies. In doing so, many different regularities in the patterns of cooperation have been found (Fehr & Schurtenberger, 2018; Henrich & Muthukrishna, 2021; Kurzban et al., 2015).

3.1.1. Common experimental approaches

By far the most common approach is to apply some variation of an economic game (of which there are several kinds). The basic idea behind these games is to approximate the social dilemmas described in the literature review (chapters 1 and 2) by providing participants with units or points that are worth real money. These games then involve participants deciding how to use or distribute these points. According to game theory, in most cases, the rational decision is to keep the maximum number of points for yourself[14]. The games can either be one-shot or iterated (played across multiple rounds with the same or different players).

In the one-shot case, three games have received the greatest attention, all of which are played between two players. In the dictator game, one individual receives a pot of money which they may solely decide on how to split with their partner (Kahneman et al., 1986). Rationally, this player has no reason to ever offer any of this pot to their partner, so this experiment is designed to assess altruistic behaviour. In a review of over 600 experiments, individuals were found to give on average 28% of the maximum donation, which also varied alongside a variety of demographic characteristics (Engel, 2011). A very similar game is the ultimatum game (Güth et al., 1982; Nowak et al., 2000). This game is identical to the dictator game, except the second player may choose to accept or reject the first players offer. If the offer is rejected, neither player receive anything. While game theory predicts that individuals should offer the lowest amount possible and their partner should accept any offer, typical offers are around 40 – 50% and offers below 30% are commonly rejected (Güth & Kocher, 2014). Additionally, there is also substantial variation in donation rates in the ultimatum and dictator games when they were applied across cultures (Henrich, Boyd, & Bowles, 2005). In the trust game, one player sends points to a second player which are then tripled, and the second player returns a chosen amount (Berg et al., 1995; Johnson & Mislin, 2011). Consistent with ultimatum games, donors typically send around 50% and partners typically return around 30% (Johnson & Mislin, 2011). The overall message of these games is clear. People do not behave rationally selfish as economic models predict, but they are not perfectly cooperative either.

One-shot economic games have also been used to study spiteful behaviour, albeit less frequently. In a classic version of the spite paradigm, the money burning game, participants were offered the opportunity to pay a portion of their earnings to remove a portion of their partner's earnings (Zizzo & Oswald, 2001). Like in the case of games designed to assess cooperation, paying to reduce your

[14] There are variations in the payoff structure of these games where game theory doesn't predict keeping the maximum number of points for yourself. Examples of these are discussed in chapter 2.

partner's earnings would not be consistent with game theory predictions. Yet, many participants chose to burn money in this experiment (Zizzo & Oswald, 2001). Later versions of the game, termed the Joy of Destruction, altered the paradigm to include the randomised destruction of money which allowed participants to hide their own spiteful decisions. In these games, up to 40% of participants chose to reduce their partner's earnings (Abbink & Herrmann, 2011; Abbink & Sadrieh, 2009).

While useful, one-shot games do not permit investigations of the dynamics of cooperation across a longer timeframe, and they are often restricted to two players. Instead, many studies use a variation of a public goods game played with multiple players across multiple rounds (Ledyard, 1995). During a public goods game, participants are granted points which they can donate to a shared pot which is multiplied and split between all players. Each round, their score (and earnings) is determined by the amount they keep from their original total plus what they earn from the pot. In the absence of other maintenance mechanisms (discussed in chapter 2), offerings usually begin around 50% but decline as the game progresses (Dawes, 1980; Dawes & Thaler, 1988; Neugebauer et al., 2009; Zelmer, 2003). In these games, participants can usually be categorised into strong co-operators, free riders (never contribute) or conditional co-operators (contribute if others are contributing) (Kurzban & Houser, 2001). Many other consistent trends have been found in such games (reviewed in, Fehr & Schurtenberger, 2018). For example, contributions usually reset to 50% if participants are assigned to a new group (Andreoni, 1988) but permitting participants to choose their partners (rather than randomly assigning them) results in higher cooperation (Gürerk et al., 2014). These games have allowed researchers to study more complex dynamics of cooperation such as reputational effects (Milinski et al., 2006) and costly punishment (Balliet, Mulder, et al., 2011; Fehr & Gachter, 2000).

3.1.2. The issue with experiments

Despite their usefulness, laboratory experiments are often criticised on several grounds. Epistemologically, many argue that by ignoring participant intentionality, it is impossible to fully understand a participant's actions (Scotland, 2012). One notable example of this within cooperation research is costly punishment. Early research assumed punishment behaviour reflected an altruistic preference for the enforcement of norms (Fowler, 2005; Gintis, 2000). However, later research revealed that punishment was often motivated by envy, not anger, at a norm violation (Pedersen et al., 2013) and given the choice, participants would rather compensate an individual who lost out, than direct punishment towards a non-co-operator (Pedersen et al., 2018). Furthermore, experiments are (necessarily) simplistic representations of reality, which in turn compromises the external or ecological validity of experimental paradigms (Bryman, 2004). This raises some concern as to the extent that economic game results may be extrapolated to the real world (Pisor et al., 2020). Other studies find little correlation between laboratory and real-world behaviour (List, 2006) or that laboratory behaviour may not reflect real-life behaviour (Benz & Meier, 2008). Seemingly minor elements of experimental design can also impact results. For example, the way the game is framed, or the amount of money participants are playing for can affect the amount participants contribute (Levitt & List, 2007). Many experimental games are also primarily comprised of student samples from WEIRD populations (Henrich, Heine, et al., 2010), which introduces further generalisability concerns. Finally, others have argued that participant confusion can bias their decision making within economic games which complicates inferences from experiments (Ferraro & Vossler, 2010). This can partially be managed by providing sufficient detail in experimental instructions to minimise the possibility of participant confusion (Ferraro & Vossler, 2010).

3.1.3. Alternatives to experiments

The two major alternatives to laboratory experiments are field-based studies and simulation models. Models are useful as they permit the assessment of the long-term evolutionary consequences of different behavioural strategies or mechanisms on the evolution of cooperation. Models usually include a replicator dynamic (Taylor & Jonker, 1978) where traits are expressed in subsequent generations in proportion to their relative fitness. In evolutionary terms, this can be thought of as differential reproduction. Within a prisoner’s dilemma (PD), the replicator dynamic will invariably spread defection or low cooperation. Therefore, many models introduce additional mechanisms to investigate how they change this dynamic. Among the earliest modelling work within cooperation was a computer tournament conducted by Axelrod and Hamilton, (1981) who allowed various behavioural strategies to compete in iterated binary PD games. They found the strategy Tit for Tat (match partners contribution) emerged as the overall most successful strategy in repeated 1-1 interactions. This provided initial evidence for the theory of direct reciprocity (Trivers, 1971). Other more recent models have addressed a substantial variety of other questions including the effect of spatial structure (Nowak & May, 1993), punishment (Boyd et al., 2010) and different payoff structures (Doebeli & Hauert, 2005a) and continue to provide unique insights difficult to obtain through other methods. While simulation models do circumvent some issues intrinsic to experiments (sample issues and their scale), they are further detached from real-life social behaviour than experiments, as all “individuals” within an agent-based simulation are merely stochastically operating computer code. However, it is precisely this simplicity that makes simulation models useful, as they permit more fine-tuned investigations of specific variables of interest than would be (feasibly) possible through other methods (Acerbi et al., 2020).

A further alternative to laboratory experiments are field-based studies. Field studies typically offer greater ecological validity than do laboratory or simulation-based studies as an individual’s cooperative behaviour is measured in an applied setting. Some argue that it is essential to study people within natural settings to fully understand their behaviour (Eisner, 2004). Field studies also allow researchers to confirm or test predictions that arise from models or experiments. Many studies within cultural evolution and cooperation have done this successfully. Payoff biased transmission was found for opening plays in the game GO (Beheim et al., 2014) and managers choice of football formation (Mesoudi, 2020) and applying cooperation science to a social dilemma regarding the production of food resulted in several practical policy suggestions including food coops and improved loan schemes (Hupper et al., 2019). Assessing trends within a field context may also be more consistent with emerging theoretical predictions regarding social learning strategies, where individuals may apply multiple strategies simultaneously or synergistically (R. Kendal et al., 2018) which varies further alongside individual differences and context (Muthukrishna et al., 2016; Rawlings et al., 2017; Vale, Flynn, et al., 2017).

Stepping beyond quantitative to qualitative methods offers several other benefits. Qualitative methods provide a richer representation of the phenomena of interest by considering the meaning behind behaviour (Braun & Clarke, 2013) and are favoured on epistemological grounds by many researchers (Eagly & Riger, 2014). In comparison, quantitative methods are argued to be less able to capture the characteristics of the day-to-day social world than qualitative methods (Howitt & Cramer, 2008). However, qualitative methods come with their own limitations. Many have argued that qualitative research is inherently more subjective, as researchers bring their biases and interpretations into their research (Hood, 2006). Because researcher-participant interactions are a more integral part of the qualitative research process, researchers may influence the responses participants give (Kvale, 1995). Epistemologically, relying on what individuals claim they do rather than what they do in practice can also be problematic (Hammersley, 1990). This is especially apparent in cooperation research, where social desirability may incentivise many people to overclaim their cooperative tendencies. Consistent with this, cooperation was around 25% higher in a hypothetical than incentivised PD game (Lönnqvist et al., 2011). Qualitative research and field studies are also inherently less controlled than experiments which can make it more difficult to investigate relationships between individual variables. Therefore, quantitative approaches may be better suited to investigate research questions that are concerned with a limited set of particular variables (Creswell & Clark, 2011). Qualitative studies also differ in the kind of generalisation they seek compared to quantitative studies[15]. Quantitative studies typically strive for a representative sample where results can be generalised to the wider population (B. Smith, 2018). Whereas one goal of qualitative research may be to identify trends relevant to specific populations which can then be compared to trends within different populations (B. Smith, 2018).

The divide between the advantages of quantitative and qualitative methods has led many to advocate for the use of mixed methods approaches that combine qualitative and quantitative methods (Creswell & Clark, 2011; Johnson et al., 2007; Schoonenboom & Johnson, 2017). There are many reasons why

[15] Generalisability as applied to quantitative studies usually takes the form of statistical-probabilistic generalisability while qualitative studies may instead strive for analytical generalisability (B. Smith, 2018).

mixed methods designs are useful. These include comparing and corresponding results between methods (triangulation), using the results from one method to inform the other (development) and evaluating different aspects of the same research question in different ways (expansion) (Greene, 2007). The quantitative and qualitative aspects of mixed methods research can happen in sequence (where the goal is building upon the findings of the previous method) or in parallel (where both methods answer the same research question) (Schoonenboom & Johnson, 2017). Across the whole thesis, the overall purpose of the mixed methods approach is triangulation, where the results obtained from the experiments can be compared with those from the field setting. However, the design of chapter 6 is motivated by an expansion approach (Greene, 2007), where quantitative and qualitative approaches are used to assess the same overarching research questions.

3.2. Methods employed in this thesis

This thesis employed a mixed-methods approach. As both cooperation and social learning research is conventionally dominated by experiments or quantitative approaches, this presents a potentially interesting gap where qualitative methods may be applied. In this section, I will outline the method I use in each empirical chapter and offer some justifications and ways that any limitations will be managed. Details of each study design are specified in the following three research chapters. The overall goal of the thesis was to investigate the extent that individuals used social learning (specifically conformity, prestige and payoff bias) in their cooperative decision making. This represents an ideal application for a mixed-methods approach as it will be possible to evaluate the extent to which laboratory and theoretical results apply to a real-life setting. Employing mixed methods helps alleviate the potential pitfalls of any given methodological approach, where, taken together, each method may compensate for the weaknesses of the other.

3.2.1. Study 1: Social Learning Strategies and Cooperative Behaviour: Evidence of Payoff Bias, but Not Prestige or Conformity, in a Social Dilemma Game

In the first empirical study (Chapter 4), I experimentally investigated the extent to which individuals used a conformity, payoff or prestige bias within their cooperative decision making. To do this, I employed a standard 4 player public goods game (Ledyard, 1995) and compared a prisoner's dilemma payoff structure with a snowdrift payoff structure (Doebeli & Hauert, 2005a). Before playing, participants completed a 10-item quiz that assessed their general understanding of social behaviour. After which, the top-scoring participant was publicly congratulated. The purpose of this quiz is to assign prestige to one of the participants, which has successfully been done in previous experiments (Brand et al., 2020; Kumru & Vesterlund, 2010). Within the public goods game, in the social condition, participants are granted access to social information which describes the donation behaviour of the other participants with specific attention drawn to; the top-scoring participant (payoff bias), the average behaviour (conformity) and the behaviour of the top scorer in the quiz (prestige bias).

The study design is novel in two major ways. Firstly, it is the first study (to my knowledge) to apply a continuous snowdrift game across multiple rounds. Many studies investigating the snowdrift game are restricted to binary decisions (Brown & Vincent, 2008; Zhong et al., 2008). Secondly, it is the first study of its kind to offer participants simultaneous access to all three of the major social learning strategies considered in the literature (prestige bias, payoff bias and conformity). Similar studies of this kind are restricted to one source of information at a time (Burton-Chellew, El Mouden, et al., 2017; Molleman et al., 2014; P. van den Berg, Molleman, & Weissing, 2015). Additionally, this experiment also compliments only one other experimental study which explicitly investigated the role of prestige or status on cooperative behaviour (Kumru & Vesterlund, 2010).

To evaluate the influence of social learning and infer which strategy (or combination of strategies) participants used, I used model comparison (widely applicable information criteria) where competing models are compared on their ability to predict out of sample (McElreath, 2020) and plotted the predicted effects and the posterior distribution of the parameter values. Allowing participants free access to social information and inferring their strategy use through model comparison offers greater ecological validity than experimentally manipulating participants access to information. Further, it also accounts for recent acknowledgements within cultural evolution that individuals often use strategies interchangeably or in combination (R. Kendal et al., 2018). Finally, the results of the experiment were used in an agent-based simulation model to evaluate the impact of group size and other parameters on the impact of social learning over time. This allowed the results to be examined beyond the experimental context which is desirable from a theoretical and external validity point of view. This represents a further novel application of Bayesian posterior distributions, where the Bayesian parameter estimates are used to predict the influence of social learning over time and selection is modelled on the posterior intercept distributions.

The experiment was coded and run using Dallinger (Dallinger, 2020) which recruited participants from Amazon's Mechanical Turk (MTurk). Though this still introduces some bias with respect to sampling (for example, by favouring those with access to a computer), it produces a more diverse sample than that of students. Additionally, economic game results from MTurk samples have been shown to correlate well with results obtained from traditional in-person samples (Rand, 2012). The sample for this experiment was 286 participants. Of those that provided demographic information, there was a median age of 32.5 years comprised of 181 men, 97 women and 1 non-binary individual. Despite equal recruitment between conditions, uneven completion rates left some groups with less complete data than others. The sample size between conditions is given below in Table 3.1. The appropriate sample size for this experiment was determined using simulated data (appendix 3.1). Models fit to different sample sizes were compared to a population model to check how closely they approximated the true parameter values.

Table 3.1. *Sample size from study 1 across the 4 experimental conditions.*

Prisoner's Dilemma		Snowdrift	
Asocial	Social	Asocial	Social
20 Groups (N = 73)	17 Groups (N = 61)	24 Groups (N = 84)	20 Groups (N = 68)

3.2.2. Study 2: Social learning does not influence spiteful behaviour: Individuals behave altruistically, but not spitefully, in an online game

The second empirical study (chapter 5) investigated spiteful behaviour, the often-forgotten kind of social behaviour (Gardner & West, 2004) and the extent that this is influenced by social learning. I employed a modified version of the Joy of Destruction game (Abbink & Sadrieh, 2009) where participants are offered the opportunity to pay a personal cost to reduce the earnings of another participant. My paradigm differed from that of typical research investigating spite as it also offered participants the opportunity to behave altruistically (by increasing their partner's score) and placed decisions on a continuum to allow differing degrees of spite or altruism. This decision arose from a concern for ecological validity, as it is artificial to assume that individuals would encounter situations where their only two options are harming another individual or doing nothing. Like chapter 4, to evaluate the impact of social learning, I used a mixture of model comparison, predictive plots and the posterior distribution of the parameter values.

For this experiment, some deception was necessary. To evaluate the impact of social learning, participants not in the control condition received fictitious information regarding the behaviour of previous participants. Specifically, the highest-scoring previous participant (copy-the-successful) or the majority of previous participants (conformity). These individuals were reported to have increased, decreased or not changed their partner's earnings when offered the same choice. Additionally, participants were led to believe that their partner was a real participant, when in fact they were playing with a bot throughout. This did avoid potential ethical concerns associated with participants reducing other individuals' earnings. Nonetheless, at the end of the experiment, participants were fully debriefed, the deception was explained to them and offered the opportunity to withdraw from the study (though, very few did).

This experiment was also coded and run using Dallinger (Dallinger, 2020) and similarly recruited from MTurk. 253 participants took part in the experiment. Of those that provided demographic information, 197 identified as male, 75 as female and 2 as non-binary. Due to uneven completion rates, there was some imbalance between experimental conditions (Table 3.2).

Table 3.2. *Experimental conditions used in study 2. Social information shown to participants varied according to its source (majority or successful) and the source's behaviour to their partner. Numbers in brackets show the sample size between conditions. The control condition not shown on this table contained 54 participants.*

Information source (Factor 1)	Source's behaviour (Factor 2)		
Successful	Reduce points (47)	No change (57)	Increase points (41)
Majority	Reduce points (53)	No change (43)	Increase points (51)

3.2.3. Study 3: Exploring social learning and sustainability in a student population: A mixed-methods investigation

The third empirical chapter (chapter 6) was comprised of a two-part study that investigated social learning in the applied cooperative domain of environmental sustainability. The purpose of these studies was to address the concerns raised by some researchers that laboratory results may not correlate with real-life behaviour (Pisor et al., 2020) and evaluate the extent to which social learning (and my experimental results) might apply to real-life cooperative behaviour. These studies deviated from their initial design following restrictions from COVID-19. Originally, this study planned to use the questionnaire to identify common sources of influence on students' sustainable behaviour which would then inform an energy intervention in Durham colleges. The proposed start date for this research coincided with the first national lockdown, and students were largely home learning in the following 2020 – 2021 academic year[16]. Consequently, this research was no longer feasible and was cancelled. Instead, the questionnaire was expanded in scope to include a behavioural measure which provided some analogy to the original field study. This was then complemented with semi-structured interviews (conducted online) which sought to expand the scope of the questionnaire and explore the themes in greater detail.

Both parts of the study were conducted with Durham University students which does introduce a concern of generalisability in terms of WEIRD samples (Henrich et al., 2010). However, many qualitative researchers argue that generalisability need not aim to produce representative samples, as instead, insight can come from comparing results and generalities from different populations (B. Smith, 2018). Practical considerations meant achieving a more conventionally representative sample was not possible. Therefore, the sample was restricted specifically to Durham University students and the scope and aims of the study were intentionally limited to evaluating trends within this population. Thus, this study does not seek to make any claims about the extent to which its findings may also apply to other populations.

The first part of this study employed a questionnaire (designed in Qualtrics) that was administered to Durham undergraduate and postgraduate students. The survey was advertised in a variety of places, including student bulletins, word of mouth and Durham Create + SONA systems (which recruit from student pools). The questionnaire included several questions designed to assess trends found within other studies into sustainable behaviour (for example, barriers to their sustainability) and in many cases adapted existing scales. This included the environmental concern scale (Minton & Rose, 1997) and general ecological behaviour (F. G. Kaiser & Biel, 2000). The questionnaire also included questions focused on conformity, payoff bias and prestige bias and asked participants to self-report the extent that this, and other, kind of information influenced their sustainable decision making. The

[16] An additional, smaller-scale study was trialled between October – December 2020 investigating hand sanitiser use in response to social prompts during the limited in-person teaching during that timeframe. This also failed due to practical issues and was abandoned following the second national UK lockdown.

questionnaire also included an experimental manipulation. Participants read short vignettes describing a fictitious (gender-neutral) individual "Alex" as either high status and promoting sustainability (prestige), noticing various benefits of a sustainable lifestyle (payoff) or as living in a community where sustainability is becoming more common (conformity). As thanks for participating, participants could enter a raffle for a chance to win one of two vouchers: An ASOS clothing voucher (less sustainable option) or a Rapanui clothing voucher (more sustainable option). This choice was then used as an outcome variable in a regression model alongside other variables and the experimental condition.

After data filtering, 182 participants were left in the dataset for analysis. Students were recruited from a range of departments, however the four most common were Psychology (N = 51), Anthropology (N = 34), Biological Sciences (N = 14) and Geography (N = 13). Of these, 30 students identified as male, 147 identified as female, 3 as non-binary and 2 withheld the information. The median age was 20 years (IQR = 3).

In the second part of the study, I used semi-structured interviews to address and expand upon the various themes of the questionnaire. The primary goal of these interviews was to provide further insight into the issues addressed by the questionnaire and evaluate the relative importance of these issues to participants. Like the questionnaire, I asked participants to comment on the barriers to their sustainability and discuss their attitude towards and understanding of sustainability. I also asked participants to comment on the influence of others on their sustainability but focused specifically on the role of conformity, payoff and prestige bias. To analyse the data, I employed thematic analysis (Braun & Clarke, 2006, 2019), as it represents a highly flexible approach for the analysis of qualitative data. I combined inductive and deductive thematic analysis, where themes from questions regarding attitudes and perceptions were generated exclusively inductively and themes addressing the influence of others were generated inductively and deductively. This was appropriate, as the questions regarding attitudes and perceptions were broader in focus and I had little preconception of the potential codes and themes. Whereas I specifically sought to evaluate the evidence for conformity, payoff and prestige bias, and codes and themes were generated with this in mind, suiting a more deductive approach.

Interviews took place online over ZOOM and were transcribed in an intelligent verbatim way, where utterances (erm, um) and erroneously repeated words were removed. This was appropriate because my interest was exclusively in the content of the interviews, rather than other conversational dynamics (Lapadat, 2000). Transcription was done initially by ai (Otter.ai, 2021) which was then manually corrected using a recording from Open Broadcaster Software. Analysis of the data was conducted in Nvivo (March 2020 version).

Participants were recruited directly from the questionnaire sample and were contacted if they had indicated their willingness on the questionnaire (though, many declined or did not respond). Therefore, all participants interviewed had also completed the questionnaire. The sample comprised of 6 Females and 1 Male (median age 21) and were all current Durham University students from a mix of departments (Table 3.3). Note that all participants were assigned pseudonyms.

Table 3.3*. Participant demographics, department and year of study (names given are pseudonyms).*

Name	**Gender**	**Age**	**Department**	**Year of study**
Zoe	Female	19	Anthropology	First
Caitlyn	Female	21	Geography	Third
Ed	Male	21	Sociology	Third
Annie	Female	20	Philosophy	Second
Evelyn	Female	22	Sociology	Third
Cassie	Female	30	Education	Fourth +
Diana	Female	20	Anthropology	First

3.3. Statistical approach

Throughout the thesis, I used Bayesian models to evaluate the research questions. The decision to employ Bayesian (rather than frequentist) statistics was both pragmatic and theoretical. Discussions regarding the issues of the overreliance on P values within science are not new. In early writings, Cohen (1995) raised concerns surrounding null hypothesis significance testing, pointing out that, despite popular belief, P does not (and cannot) tell you the probability of a given hypothesis, only that the data is unlikely under a null hypothesis. He expressed additional concerns regarding the validity of the "null" hypothesis, as 0 relationship between two variables is unlikely in practice. Many years on, these issues remain relevant. P values are still regularly misunderstood by researchers and the null hypothesis still represents an effect of 0 (Lakens, 2021). Researchers from many fields have pointed out the issues with p trending towards 0 in larger samples (Lin et al., 2013) or failing to report (or consider) the effect size and practical significance of statistically significant effects (Sullivan & Feinn, 2012).

Among others, a proposed solution is to employ Bayesian approaches which place greater focus on effect sizes, uncertainty and the practical significance of identified effects (Kruschke, 2015; McElreath, 2020). Though, it is true that when they are applied properly, frequentist and Bayesian models usually produce the same qualitative conclusion (Lakens, 2021). Besides the theoretical benefits of Bayesian models, they also offered a key practical application for this thesis. Bayesian models produce posterior distributions of possible parameters that could have resulted in the data observed, which in turn can be used to produce a distribution of predictions from the model. This is useful in simulation (agent-based) models, where it is possible to use experimental results to inform the behaviour of the simulation and therefore investigate experimental patterns over a longer time frame. It is for these reasons, that Bayesian approaches were used in this thesis. The precise models used in each case are described in the respective chapter where the mathematical notation of the model is also given. To evaluate the effects found in each study, I combine inferences from model comparison (widely applicable information criteria), plotting parameters and counterfactual predictive plots.

3.4. Conclusion

Cooperation is a wide and diverse topic of research, as is cultural evolution and the study of social learning, which bridges many disciplinary and methodological boundaries. There are a huge number of different approaches that have been taken in order to study them. This thesis seeks to evaluate its research questions in the widest way possible by applying a mixed-methods approach combining experiments with simulation modelling and interviews. By doing so, each method may cover the weakness of the other and the evidence for social learning within social dilemmas can be evaluated in the broadest and most ecologically valid way possible. Employing a field-based outcome and interviews moves beyond typical cooperation research which is primarily dominated by experiments and agent-based (or mathematical) models. Additionally, the approach to study spiteful behaviour is novel and builds upon previous studies by considering a wider range of possible behaviours. Finally, assessing the evidence for social learning strategies in real-life cooperative behaviour will offer further insights into the extent to which theoretical and experimental predictions apply outside of these limited contexts.

Chapter 4: Social learning strategies and cooperative behaviour: Evidence of Payoff Bias, but not Prestige or Conformity, in a social dilemma game

4.1 Abstract

Human cooperation, occurring without reciprocation and between unrelated individuals in large populations, represents an evolutionary puzzle. One potential explanation is that cooperative behaviour may be transmitted between individuals via social learning. Using an online social dilemma experiment, we find evidence that participants' contributions were more consistent with payoff-biased transmission than prestige-biased transmission or conformity. We also found some evidence for lower cooperation (i) when exposed to social information about peer cooperation levels than without such information, and (ii) in the prisoners' dilemma game compared to the snowdrift game. A simulation model established that the observed cooperation was more likely to be caused by participants' general propensity to cooperate than by the effect of social learning strategies employed within the experiment, but that this cooperative propensity could be reduced through selection. Overall, our results support previous experimental evidence indicating the role of payoff-biased transmission in explaining cooperative behaviour but we find that this effect is small and is overwhelmed by participants' general propensity for cooperation.

4.2 Introduction

The spread of behaviour that benefits others is difficult to explain through natural selection, as such behaviour risks exploitation from others (Binmore, 2007b). Scenarios where prosocial behaviours can be exploited by others are termed social dilemmas (Dawes, 1980). Classic mechanisms to maintain cooperation include kin selection (Hamilton, 1964), punishment of non-co-operators (Balliet, Mulder, et al., 2011) and reciprocity (Trivers, 1971). Given this, human cooperation is especially surprising because it occurs between unrelated individuals and is often unreciprocated (Henrich & Muthukrishna, 2021; Kurzban et al., 2015). Laboratory studies (usually from WEIRD samples, but see (Henrich et al. 2005) also show that individuals often cooperate at higher levels than would be predicted by game theory (Fehr & Schurtenberger, 2018).

Forms of cooperation may be culturally transmitted within or across social groups through social learning (Richerson et al. 2016; D. Smith 2020). Through social learning, individuals acquire traits or information by observing or interacting with other individuals or the products of their behaviour (Heyes, 1994). Social learning allows individuals to obtain adaptive traits that are difficult to acquire asocially but can also result in the spread of outdated or maladaptive information (Mesoudi, 2009; Rieucau & Giraldeau, 2011). For this reason, complete reliance on social learning is unlikely to be adaptive (Boyd & Richerson, 1995; Schlag, 1996b). Instead, scholars have suggested that selection should favour strategic use of social learning via strategies that influence when, what and from whom individuals socially learn (Laland, 2004; Morgan et al., 2012). Three strategies have received particular attention: payoff bias (copy traits that yield a high payoff; (Schlag 1996); conformity (disproportionate propensity to copy common traits; Henrich and Boyd 1998); and prestige bias (copy individuals of high status; (Henrich and Gil-White 2001). Kendal et al., (2018) reviews evidence for contexts in which these strategies are used, either individually or in combination.

Given that both cooperation and social learning are thought to underpin the massive habitat expansion and the evolution of complex cultural systems characteristic of our species (Henrich, 2018; Powers et al., 2016), it is perhaps surprising that relatively few studies have addressed how they interact. Conformity may be able to sustain cooperation when combined with punishment (Lewis & Ekart, 2017) or when cooperation is already common (Carpenter, 2004). In the lab, participants conformed to an external group's donations (Bardsley & Sausgruber, 2005) or cooperated with a previously cooperative partner (Romano & Balliet, 2017) but direct reciprocity was an overall stronger influence on behaviour. Henrich & Gil-White (2001) suggest that through prestige, followers grant voluntary deference towards leaders in exchange for learning opportunities. This could incentivise copying of cooperation among followers and increased cooperation from leaders (Henrich et al., 2015). Models suggest that prestige can maintain cooperation in a larger range of scenarios than other social learning strategies (Henrich et al., 2015; Lehmann & Feldman, 2008; Molleman, Quiñones, et al., 2013). In the lab, participants have exhibited a bias to copy large contributions made by leaders (Gächter et al., 2012). Furthermore, prosocial leaders (measured by a questionnaire) elicited greater cooperation from their group than selfish leaders and used punishment less than traditional peer sanctioning groups (Harrell & Simpson, 2016). Experimental and ethnographic studies suggest that leader fairness and charisma can positively affect cooperation (De Cremer and Van Knippenberg 2002). There is evidence in strictly hierarchical institutions that team performance and information flow is correlated with the degree of informal prestige conferred upon leaders (Offord et al., 2016, 2019) although, contrary to common marketing strategy, there is also evidence that real-life cooperative behaviours are not highly influenced by celebrity endorsement (John et al., 2019; N. T. Wood & Herbst, 2007). Formal status or rank has received little attention, although Kumru & Vesterlund (2010) found participants labelled with stars (from a superior quiz performance) were copied more than those without stars.

Because payoff-biased social learning results in the adoption of traits proportional to their relative fitness, as formalised in the replicator equation (Schlag, 1999; Taylor & Jonker, 1978), it would be expected to spread selfish behaviour. (Grujić and Lenaerts 2020) experimental study showed that participants exhibited a bias to copy their more successful neighbours and reduce their cooperative contributions to a public good. Experimental evidence suggests that participants are more likely to exhibit a payoff bias than conformity in a cooperation game and reduce their contributions (Burton-Chellew, El Mouden, et al., 2017; Molleman et al., 2014) and also decrease their contributions when reminded how their behaviour was benefiting others (Burton-Chellew et al., 2015). A recent analysis on 237 PGGs also showed that declines in contributions were most consistent with improving personal payoffs (Burton-Chellew & West, 2021). Furthermore, cooperation was also higher when participants had no information on the behaviour of their group mates (Neugebauer et al., 2009). Payoff biased learning may not be detrimental for cooperation in all cases, for example when defection is less rewarding (Szolnoki & Chen, 2018) or when group migration and punishment is possible (Gurek et al., 2006). Generally, it results in the decline of cooperative behaviour and is the kind of information preferentially attended to.

While strategic defection can maximise payoffs, the patterns of results point towards payoff biased social learning being the preferred strategy adopted by participants in social dilemmas. While conformity can increase cooperation in some contexts, it appears to be the weakest cue when compared with other social learning strategies (Burton-Chellew, El Mouden, et al., 2017; Molleman et al., 2014; Molleman, Quiñones, et al., 2013). Prestige (specifically, high status) biased social learning is comparatively understudied in cooperative dilemmas but is predicted to sustain cooperation in a wide array of circumstances (Henrich et al., 2015; Lehmann & Feldman, 2008). Because no experimental study has considered all three strategies simultaneously in a cooperative context, this is the primary aim of our study.

The evolution of cooperation can also be affected by the payoff structure of the social dilemma. Typically, cooperation games assume a prisoner's dilemma (PD) payoff structure (see Table 4.1), where game theory always predicts defection as the rational choice (Rapoport, 1974). An alternative is the snowdrift game (SD, see Table 4.1 for payoff structure), also sometimes referred to as the chicken game or the hawk-dove game (Doebeli & Hauert, 2005a). Whereas models based on the PD predict defection as the evolutionarily stable strategy (Axelrod & Hamilton, 1981), SD games predict stable populations of both cooperation and defection (Kun et al., 2006; Souza et al., 2009). This is because, in the SD game, exploited co-operators still outperform exploited defectors and so cooperation is favoured when defection is common. The production of enzymes in the environment by yeast and bacteria equates to a SD game as enzyme producers benefit from their enzyme production as much as defectors. As predicted, while many cells defect by abstaining from enzyme production and freeride on neighbouring cells, the production of enzymes is not extinguished (Gore, Youk, and Van Oudenaarden 2009; R. P. Smith et al. 2019).

There are fewer experimental studies using the SD game than the PD game, perhaps because the evolution of cooperation is a harder problem in the latter. Nonetheless, both scenarios can be seen in the real world. For example, climate issues are commonly seen as a PD game or dilemmas of collective action (Keohane & Victor, 2016), whereas scenarios like constructing communal flood defences or watching for predators are more akin to SD games. One experimental study comparing an iterated binary PD with a SD game found higher cooperation in the SD game (Kümmerli et al., 2007). Similar patterns have also been found in other experiments, often using one-shot binary decision games (Hilbig et al., 2018; Reed et al., 2018; Su et al., 2018). Payoff structure also affects the spatial patterns by which cooperation is predicted to evolve, where localised clusters and dendritic spines of co-operators form in models of PD and SD games, respectively (Doebeli and Hauert 2005)

Despite these patterns in findings, comparatively little is known about the dynamics of the SD game compared with the PD. While there are many examples of PD models which consider cooperation on a continuum (Roberts and Sherratt 1998), few have considered SD games along these lines (Brown & Vincent, 2008; Mcnamara et al., 2008; Sasaki & Okada, 2015; Zhong et al., 2008). Exact payoff structures vary slightly, but they each follow the characteristic hierarchy shown in Table 4.1 and described in (Doebeli and Hauert 2005). Typical findings in such models are a convergence towards contributions of around 50%. No experiments have considered iterated continuous SD games in a group context or alongside social learning, so addressing this limitation is the second aim of this study.

Table 4.1. *Payoffs associated with cooperation (C) and defection (D) depending on the behaviour of a partner between a PD and SD game adapted from* (Doebeli and Hauert 2005). *Both tables show benefit (b) compared to the costs (c). Payoff rankings are DC > CC > DD > CD for the PD and DC > CC > CD > DD for the SD game.*

	Prisoner's dilemma		Snowdrift	
	C	D	C	D
Payoff to C	*b – c*	-*c*	*b – c/2*	*b - c*
Payoff to D	*b*	0	*b*	0

4.2.1. Research questions

Our experiment addresses several key gaps in previous research. Rather than forcing participants to adopt a particular social learning strategy across experimental conditions, we adopt a more naturalistic approach by permitting participants free access to the information required for all three (prestige, conformity, and payoff bias) of the major social learning strategies in a cooperative game. We then use statistical models to infer which social learning strategies were used. Secondly, we compare both the PD and SD games across multiple rounds, which allows a comparison of cooperation rates and social learning strategy use between games beyond a one-shot context. Further, decisions are extended beyond a binary choice to allow participants to express differing degrees of cooperation. To this end, our experiment employs a between participants 2x2 factorial design with a PD and SD condition alongside asocial (no access to social information) and social learning (access to social information) conditions. This experiment addressed 4 research questions (RQ).

1) **How do social learning strategies influence cooperative behaviour?**

   1a) **Which social learning strategies, if any, do participants use? -** We predict that payoff learning will have the strongest influence on cooperative behaviour, followed by prestige and then conformity.

   1b) **Are the patterns of social learning strategies consistent across the PD and SD game?** – We make no predictions over the direction of the interaction between social learning strategy use and payoff structure.

2) **What effect does payoff structure have on cooperative behaviour? –** We predict higher cooperation in the SD game than the PD.
3) **What effect does access to social information have on cooperative behaviour?** We predict lower cooperation with access to social information than when individuals make decisions asocially because we expect a payoff bias to decrease cooperation.

4.3. Methods

4.3.1. Design

The experiment involved four conditions across participants for a between participants 2x2 factorial design. Factor one was the social information condition (Asocial vs Social) which manipulated whether the participants had access to social information (see below). Factor two was the payoff structure (Prisoner's dilemma vs Snowdrift). We used post-hoc model selection to infer which social learning strategies had been used.

4.3.2. Materials and procedure

The experiment was executed using the experimental automation platform Dallinger (Dallinger, 2020) which recruits participants via Amazon's Mechanical Turk. Upon arrival into the virtual environment, participants were assigned to a group and a unique numerical participant ID was generated for them. Once the group contained four participants, the experiment began. It was split into two parts and all participants completed both parts.

Participants first completed a ten-item quiz containing a variety of questions assessing their understanding of expected behaviour in social situations to act as a proxy for prestige (see SI 1). Each question had three possible answers with one (pre-determined) correct answer. At the end of the quiz, a public congratulation was displayed on screen for the participant who gave the most correct answers. For the rest of the experiment, this participant's ID was displayed surrounded by stars (*)

and participants were aware this identified the top scoring participant. The questions and scoring were identical across experimental conditions and at no point were participant's actual scores revealed.

Following the quiz, all four participants took part in a six-round public goods game (PGG) with either a PD or SD payoff structure, designed in accordance with typical PGGs found in the literature (Fischbacher et al., 2001). Participants did not know ahead of time, how many rounds they would be playing, but received detailed instructions on how the game worked (see SI 2) and that they would receive a bonus payment depending on their score. In the PD, in each round, participants were granted 10 points and could then decide how much of this to donate to a pot. The pot was then doubled and split evenly between all players. The points received from the pot were then added to what the participant had kept for themselves which formed their total score for that round.

The SD game had the following modification: if the donations to the pot were less than 10, all participants received nothing for that round (including losing whatever points they had kept for themselves). This was motivated both by precedent and a theoretical basis as SD models have previously employed similar payoff structures (Souza et al., 2009). Moreover, the snowdrift game requires that defection offers the best payoff against a co-operator, but cooperating is favourable to defection against another defector (Doebeli & Hauert, 2005a). In a real-life environment, this implies the public good is unreachable without some minimum investment. Therefore, a value of 10 is chosen as this allowed a single participant to meet the necessary threshold in a single round.

In the asocial condition, participants received no information about their groupmates' behaviours and only received feedback on their own earnings from the round (calculated based on the amount received from the pot and what they kept). In the social information condition, at the end of each round, participants viewed a table showing each group member's ID, their contribution to the pot for that round, their cumulative score and the average donation across the group. The participant with the most cumulative points was labelled as the current leader with text beside their ID (see SI 2). The participant who scored the highest in the quiz was also labelled with stars (*) around their name. This information could be used if participants engaged in particular social learning strategies: the average donation for conformity; the identification of the PGG leader for a payoff bias; and the identity of the individual who scored the highest in the quiz as a proxy for prestige.

Once the PGG was completed, participants were debriefed, and basic demographic information was collected. Participants also answered two short free text questions to explain their decisions in the experiment (see appendix 4.1). Participants were also asked to rate their understanding of the game to ensure participants had understood the protocol (Ferraro & Vossler, 2010).

4.3.3. Participants

Participants were recruited using Amazon's Mechanical Turk and completed the experiment online. After filtering out groups with missing data (exclusion criteria is described below) this left 286 participants in the dataset. Of those participants, for whom demographic information was available, the median age was 32.5 years (IQR = 13 years) with 181 men, 97 women, and 1 non-binary individual. Participants self-identified as White (195), Black (23), Asian (21), Hispanic (11), other (12) or did not report this information (21). The experiment took between 5 and 10 minutes to complete. Participants earned a minimum of $1 for completing the experiment but could earn a further maximum bonus of $3 dependent upon their cumulative points score. On average, participants earned a bonus of $1.62, resulting in them earning above US federal minimum wage. Despite equal recruitment across conditions, there were unequal completion rates (see Table 4.2). The required sample size was determined using simulated data (code available at osf.io/vx78c)

Table 4.2 *Distribution of participants and groups across experimental conditions.*

Prisoner's Dilemma		Snowdrift	
Asocial	Social	Asocial	Social
20 Groups (N = 73)	17 Groups (N = 61)	24 Groups (N = 84)	20 Groups (N = 68)

4.3.3. Data analysis

All analyses were pre-registered on open science (osf.io/vx78c), though minor deviations to this are explained below. Analyses were performed in R studio version 4.0.0 (R Core Team, 2021) using the packages *tidyverse*, *ggplot2* and *brms* (Bürkner, 2017). Though the use of *rethinking* was pre-registered, *brms* creates equivalent models and both packages use MCMC (Markov chain Monte Carlo) to construct the posterior distribution. Groups were excluded from analysis if either; the group had fewer than 3 participants complete the experiment (to avoid 1:1 return from investment and collapse of the social dilemma (R. M. Dawes, 1980)) or the individual who scored the highest in the quiz had dropped out. This also deviated from pre-registration, where any groups of less than 4 were planned to be dropped. This was necessary to maintain sufficient power because far more groups than anticipated were incomplete. The analysis excluded the first round's donation behaviour, as participants in the social information condition had not viewed any social information by this point. Finally, it was pre-registered that only the social condition would be used in models evaluating social learning strategies. Adding the interaction with the binary variable social meant it was not necessary to filter the data in this way.

Cooperation (whole number of points donated between 0 and 10) was treated as a categorical variable for the purposes of the models, so Bayesian ordered probit regressions were fit to the data. An ordinal outcome was appropriate for cooperation, as it is not truly a continuous measure and using a cumulative log odds link function permits non-linearity between the levels of the variable. Bayesian models were fit using the *brms* package (Bürkner, 2017) and the posterior distribution constructed using MCMC and main model results validated in JAGS. All models were fit with 4 chains of 1000 warmup samples and 5500 samples for inference, using weakly regularising priors. Model diagnostics indicated good parameter identification and model convergence (Rhat values between 1.00 and 1.01 and lowest bulk effective sample size >2500).

Multiple models were fit to evaluate each research question, each containing different combinations of predictor variables. Though each model differed in terms of additional predictor variables (described in the respective sections), all models contained a random effect of participant to account for autocorrelation between repeat observations and variability between participants. The models used to infer the effect of social learning strategies on the degree of cooperation also contained binary variable fixed effects of game structure (snowdrift), social information condition (social) and being prestigious (prestigious participant, see appendix 4.3) to act as control variables. For social, the integer 1 indicated it was the social information condition, for snowdrift the integer 1 indicated it was the snowdrift condition and prestigious participant the integer 1 indicated they were the prestigious participant.

The structure of the Prestige+Conformity+Payoff model is shown below, where Cooperation is predicted by a vector of probabilities *p* and each response value *k* is linked to an intercept parameter a_k, with additional deviation from participant level effects ($\varepsilon_{participant}$) and slopes for each of the possible predictors (e.g., payoff bias, prestige bias and conformity). This produced an estimate of the cumulative log odds for all values of cooperation.

$$Cooperation \sim Categorical(\boldsymbol{p})$$

$$p_k = \begin{cases} q_k, & k = 1 \\ q_k - q_{k-1}, & k > 1 \end{cases}$$

$$logit(q_k) = \begin{cases} \alpha_k - \phi, & 1 \leq k \leq 10 \\ Inf, & k = 11 \end{cases}$$

$$\phi = \beta_1 * snowdrift + \beta_2 * \ prestigious\ participant + \ S * social + \ \varepsilon_{participant}$$

$$S = (\beta_3 + \ \beta_4 * prestige\ bias + \ \beta_5 * payoff\ bias + \beta_6 * conformity)$$

$$\alpha_{1:10} \sim Normal(0, 1.5)$$

$$\beta_{1:6} \sim Normal(0,1)$$

$$\varepsilon_{1:286} \sim Normal(0, \sigma)$$

$$\sigma \sim Exponential(1)$$

The predictive ability of competing models was evaluated using Widely Applicable Information Criteria (WAIC) which is computed from the log likelihoods of models and a parameter penalty to ensure predictive ability is balanced against the risks of overfitting. The calculated value indicates predicted out-of-sample deviance, where lower values indicate lower deviance (and thus, better fit). Following model evaluation, model predictions were plotted to visualise the results. Note that all model predictions were generated using the average of the participant effects. For further discussions on this approach, see (Kruschke, 2015; McElreath, 2020).

4.3.4. Simulation model

We simulated the long-term consequences of participant behaviour, as estimated by the Bayesian model (see appendix 4.4 for full details of the simulation). In the basic horizontal-transmission model, agents played repeated PD games. Like the experiment, at each round agents donated between 0 and 10 units and received a payoff according to their donation and the total in the public good. Each agent's contribution for the next round, was calculated by sampling from the posterior of the Bayesian model that includes all three social learning strategies. Additionally, each agent had an intercept parameter, which can be thought of as their baseline propensity towards cooperation, independent of their social learning strategy use, drawn from the posterior distribution of intercepts estimated from the experiment.

In addition to matching the simulated conditions to those of the experiment, we also examined dynamics across different group sizes and varied agent intercept values to establish their influence on dynamics over that of the social learning strategies estimated from the experiment. We also introduced selection to examine what happens if, consistent with the replicator equation (Taylor & Jonker, 1978), the distribution of agent intercept values changes each timestep in proportion to the payoffs received by the agents.

4.4. Results

Figure 4.1 shows the mean cooperation rates from rounds 2 – 6 for the four experimental conditions. Generally, mean cooperation was around 6 points at round 2 and showed little change across subsequent rounds. This suggests that overall cooperation rates were relatively consistent throughout the experiment. Participants also generally indicated a good understanding of how the game worked (rated out of 10: median = 8, IQR = 3).

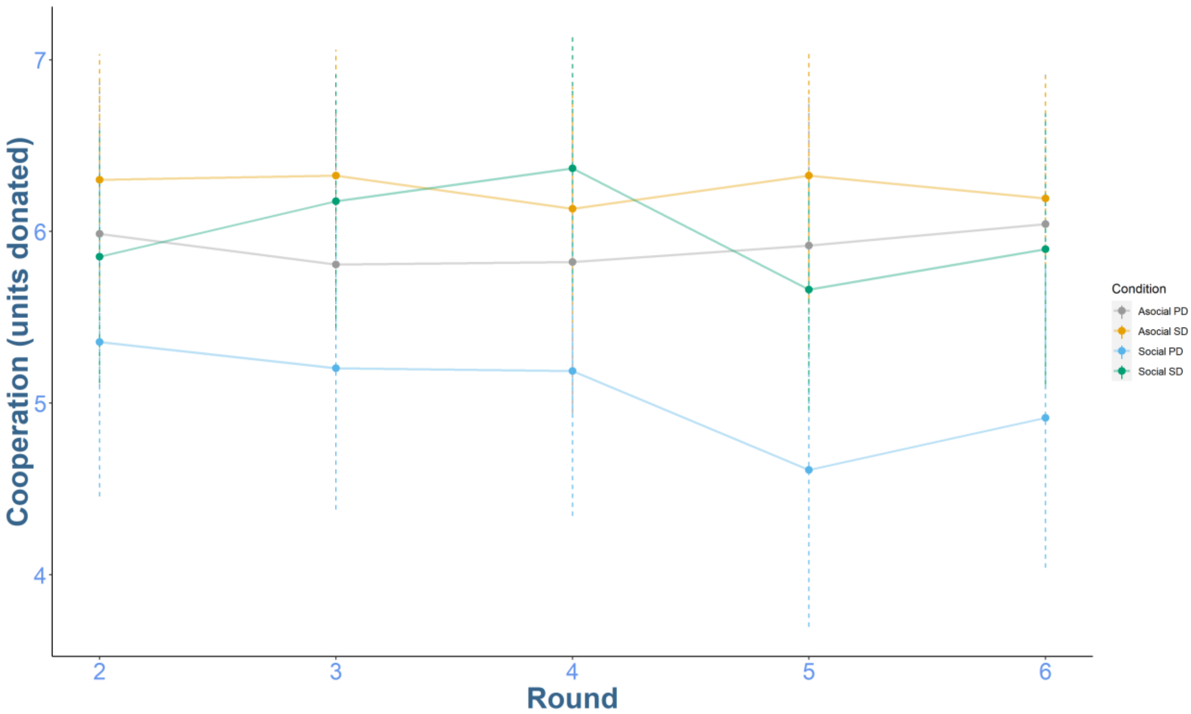


Figure 4.1. *Cooperation across rounds 2-6 between all combinations of experimental conditions. Solid points show mean cooperation and dashed lines show the mean standard error.*

Although there appears to be little variation between rounds, economic games commonly find declines in cooperation across rounds (Neugebauer et al., 2009; Zelmer, 2003). Therefore, it may still be necessary to control for variation between rounds. Two competing models were compared, one which ignored round ("No round") and another which added a varying intercept for round ("Round"). No round had a WAIC score of 5502.3 (SE = 76.9, weight = 0.73) and Round a WAIC score of 5504.3 (SE = 77.0, weight = 0.27), indicating no improvement in out-of-sample predictive ability by varying intercepts by round. The results were similar when round was included as a continuous linear predictor (No round; WAIC = 5501.4, SE = 76.9, weight = 0.71, Round; WAIC = 5504.2, SE = 77.0, weight = 0.29). Therefore, all further models excluded the effect of round.

4.4.1 Which social learning strategies, if any, do participants use? (RQ 1a)

Some modelling concerns needed to be addressed before answering this question. Data from the asocial condition was retained in the model for analysis to ensure that parameter estimates for the effects of payoff structure can be evaluated across the social and asocial condition. However, data from the asocial condition cannot be used to estimate the social learning parameters because participants did not view any social information. To address this, we modelled the interaction of the three social learning strategy parameters with the social information condition: the predicted effect is always 0 if the data come from the asocial condition.

A second concern is that, for participants who are either prestigious or have the highest payoff, the prestige and payoff social information is not strictly social as it refers to their own previous behaviour. To address this, the model used binary variables to exclude each participant from using social learning strategy data about themselves to construct the social learning parameter estimates. Specifically, prestige interacted with a binary variable where the integer one indicates they are not the prestigious individual. Payoff interacts with a binary variable where the integer one indicates they are not currently the highest earner. The result is parameter estimations occur only for cases where the slopes are not inflated by one's own behaviour.

The conformity information presented to participants (average group behaviour) included their own behaviour, but not exclusively. While reconstructing this variable to exclude their own behaviour would correct for this issue, this introduces an inconsistency between the modelled variable and the information participants were presented with in the experiment. Therefore, the proceeding analysis was repeated for uncorrected (includes their behaviour) and corrected (excludes each participant's own behaviour) conformity information. The main text details the uncorrected analysis while

Appendix 4.2 shows the main model predictions with the corrected variable and the difference in estimated parameters. Qualitatively, the primary conclusions do not differ from one another.

Eight different models were fit to the data, consistent with the constraints described above, covering all possible combinations of the social learning parameters. The compared models ranged from an asocial model containing only the control variables, to a model containing all the social learning strategies (Prestige+Conformity+Payoff). The WAIC values and associated model weights are displayed in Table 4.3.

Table 4.3. *WAIC values and model weights for models evaluating the impact of social learning strategies. Standard error difference shows the standard error in the difference between each model and the model with the lowest WAIC value. All social learning parameters interact with the social information condition. Model names indicate which social learning strategies are included. All models include the parameters for game structure and social information condition.*

Model	*WAIC*	*Standard error difference*	*Weight*
Prestige + Payoff	5499.5	0	0.34
Payoff	5499.9	1.3	0.27
Asocial	5501.2	3.3	0.14
Prestige	5501.2	3.0	0.14
Payoff + Conformity	5503.6	1.3	0.04
Prestige + Conformity + Payoff	5503.7	0.5	0.04
Prestige + Conformity	5506.6	3.1	0.01
Conformity	5506.7	3.3	0.01

The pattern of WAIC scores do not provide conclusive support for any particular social learning strategy. Overall, the strongest evidence is for payoff bias as the two top models which are favoured over the asocial model and have a combined weight of 0.61, include payoff bias. Conversely, the four models which include conformity have the lowest overall model weights (0.10), indicating models that include conformity are overfit compared to the asocial model. There appears to be a small effect associated with a prestige bias, as adding prestige to a model containing payoff does slightly improve its out-of-sample predictive ability. However, prestige alone is not favoured over an asocial model, which suggests that it is primarily payoff that is improving the model fit. Additionally, the asocial model is (modestly) favoured over those which do not contain a payoff bias or contain conformity. This includes the everything model which despite containing payoff, is penalised by WAIC for including conformity and prestige. This further suggests that conformity and prestige bias are overfit compared with payoff bias.

Parameter estimates (figure 4.2) and model predictions (figure 4.3) from the Prestige+Conformity+Payoff model are displayed in the plots below. Figure 4.3 is split between the three social learning strategies and predictions are generated for increases in the respective social learning information while holding all other variables constant. The slope for payoff is positive which indicates that, generally, participants' behaviour aligned with the direction (increase/decrease) of this social information. The slopes for prestige and conformity are weakly positive but have wider prediction intervals and the parameter estimates overlap 0. This, combined with the distribution of model weights, suggests that out of the three social learning strategies, a payoff bias shows the strongest influence on participant cooperation and therefore the changes in cooperation observed are most consistent with a payoff bias.

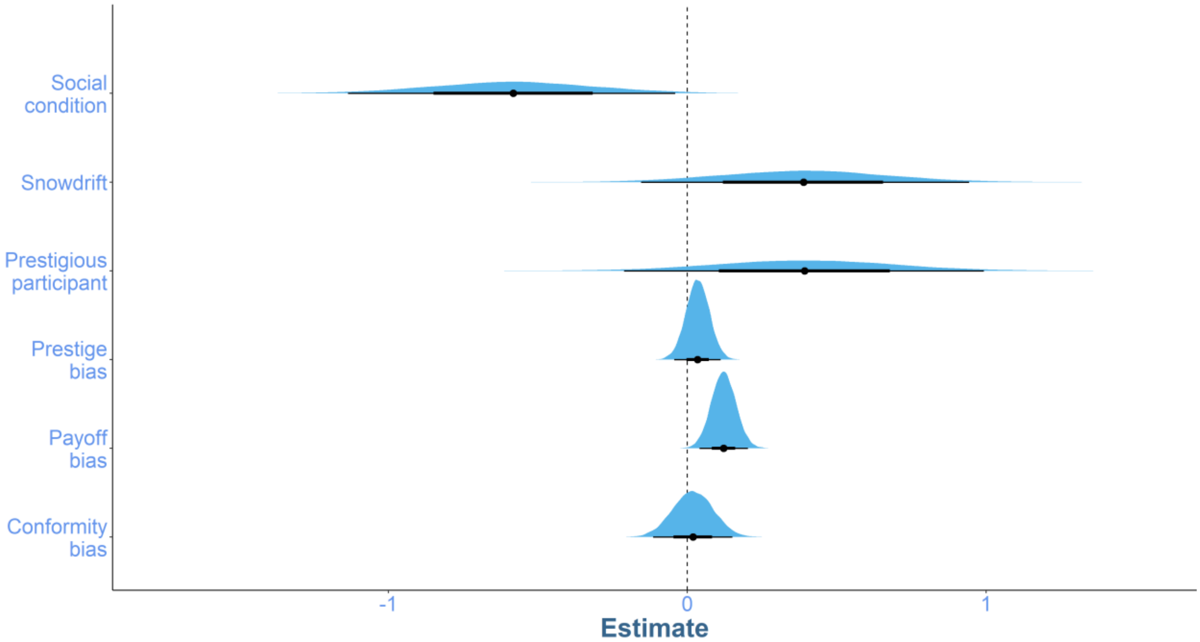


Figure 4.2. *Mean parameter estimates and 95% prediction intervals for all main effects in the Prestige+Conformity+Payoff model for the six main effects.*

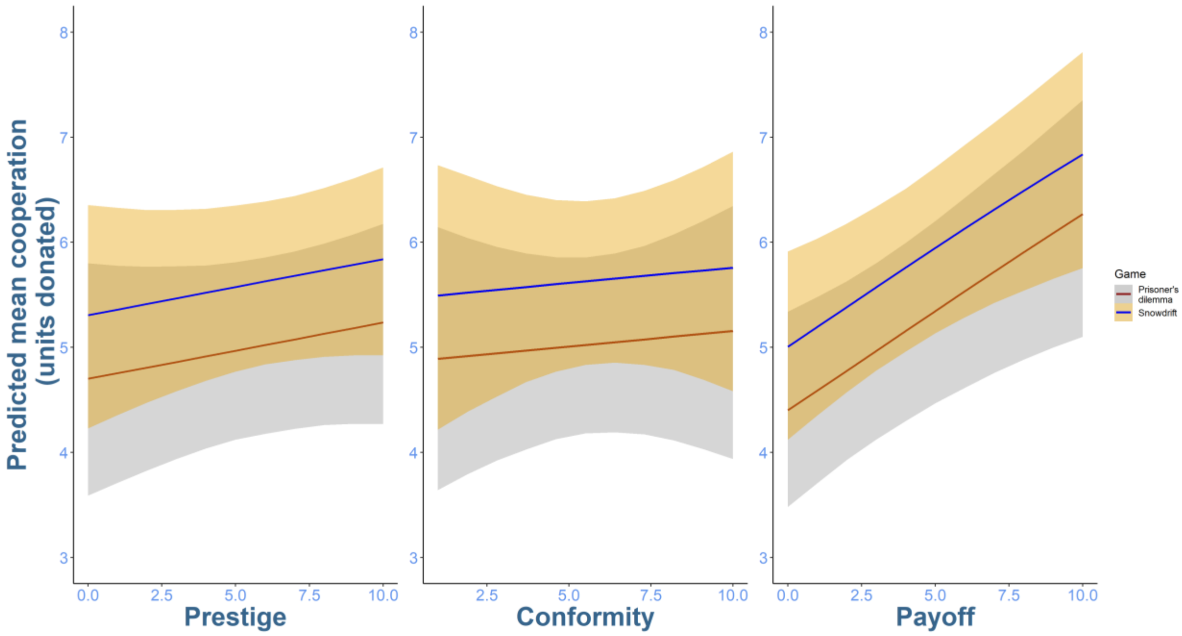


Figure 4.3. *Model predictions for Prestige, Conformity and Payoff varying the three social learning strategies and holding the other variables constant and setting the other social information at 0. Shaded regions represent the 95% prediction intervals for the* mean. *Prestige left; conformity middle; payoff; right.*

4.4.2. Are the patterns of social learning strategies consistent across the PD and SD game? (RQ 1b)

To evaluate any differences in social learning strategy use between PD and SD games (research question 1b), the Prestige + Payoff + Conformity model was compared to a model where the social learning parameters also interacted with game structure. This allowed the model to estimate different slopes for the social learning parameters between the PD and the SD game. This did not improve model fit (Everything: WAIC =5503.7; se = 77.3; weight = 0.77, Interaction: WAIC =5506.2; se = 77.6; weight = 0.23), indicating that social learning strategy use, or the influence of any social learning strategy, did not differ between the PD and SD games.

4.4.3. Evaluating the experimental conditions (RQ 2 & 3)

To evaluate the effects of game structure and the social condition (research questions 2 and 3), model comparisons were run between the Prestige+Conformity+Payoff model and models that dropped different combinations of binary variables pertaining to game structure and social condition or allowed them to interact. This means these effects can be evaluated while controlling for social learning strategy use and remain comparable to the models presented above. Every model also retained the control for being the prestigious participant. The WAIC values and associated model weights are shown in Table 4.4.

Table 4.4. *WAIC values and model weights for models evaluating the impact of binary experimental condition variables. Standard error difference shows the standard error in the difference between each model and the model with the lowest WAIC value.Model names indicate which predictor variables are included in addition to the social learning strategy parameters. The Prestige+Conformity+Payoff model is the full model containing the social learning strategies and the parameters for game structure and social information condition.*

Model	*WAIC*	*Standard error difference*	*Weight*
Social information condition only	5502.3	0	0.31
Social information condition * Game structure	5502.4	1.7	0.29
Game structure only	5503.6	1.1	0.16
Prestige + Conformity + Payoff	5503.7	0.7	0.15
No binary variables	5505.0	0.9	0.08

Overall, there is no clear distinction between any of the models. It is therefore unclear whether including either (or both) predictors (or their interaction) benefits out-of-sample model fit or not, though both top models contain social condition (combined weight 0.60).

Figure 4.4 shows model predictions generated from the interaction model. There is some indication that cooperation was lower in the social condition than the asocial condition and (to a lesser degree) higher in the SD than the PD game (Social = -0.58, 95% PI = -0.04; -1.14, Snowdrift = 0.39, 95% PI = -0.15; 0.94). Nonetheless, our predictions for research questions 2 and 3 are not clearly supported by the data.

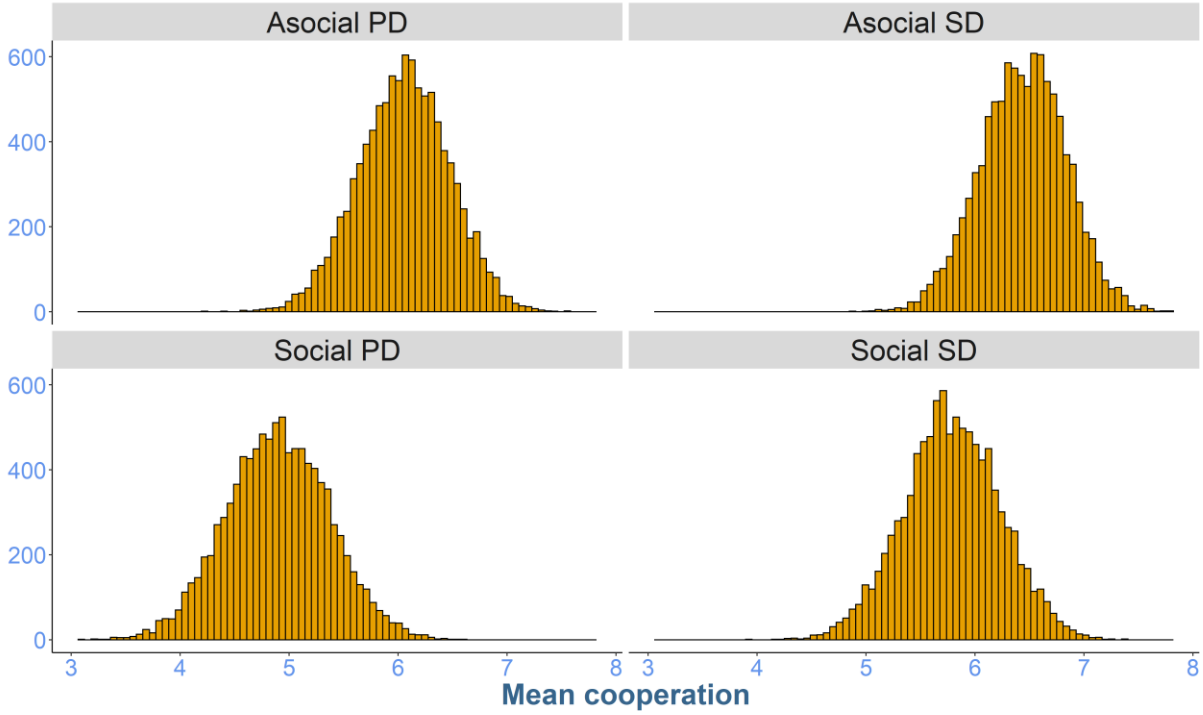


Figure 4.4. *10000 predicted values of mean cooperation across all experimental conditions generated from the Social condition*Game structure (interaction) model. Sample means and 95% PI are Asocial PD (M = 6.05; 95% PI = 5.23, 6.85), Asocial SD (M = 6.44; 95% PI = 5.65, 7.20), Social PD (M = 4.89; 95% PI = 3.96, 5.80, Social SD (M = 5.79; 95% PI = 4.91, 6.64).*

4.4.4. Simulation Model Dynamics

We used a simulation model to evaluate the longer-term consequences of the patterns of behaviour observed in this experiment. This permits the predictions from the Bayesian model (and the role of social learning) to be investigated for larger group sizes and under selection. This model samples from the Bayesian posterior estimates from the Prestige+Conformity+Payoff model to establish each agent's intercept propensity for cooperation and the influence of the social learning strategies (taking into account that the simulation model considers the PD and social condition). Note that the social learning strategies are assumed to operate non-independently of one another. Figure 4.5 shows that for the basic horizontal-transmission simulation, mean cooperation quickly stabilised to a relatively steady state at around a contribution of 5.7, indicating that social learning strategy use is not predicted to cause long-term change in the frequency of cooperation in a population. By comparison, Figure 4.6 shows that if we force agents to adopt a particular intercept propensity for cooperation (high, low), cooperation stabilises at different levels. Thus, over a long timeframe and provided participants continue to behave on average as they did in the experiment, cooperation levels are far more strongly affected by the intercept propensity for cooperation than by the effects of social learning strategies. We found that group size did not affect these qualitative findings (see appendix 4.4).

We introduced selection and small random mutation on the intercept propensity for cooperation by assuming intercept values in one round are represented in the next round in proportion to payoffs earned and then altered by a small amount by sampling from a normal distribution around the inherited intercept value. This simulation can either represent selection and mutation across biological generations, or modification of an individual's propensity for cooperation over time within a generation.

We found that cooperation declined as agents with small intercepts contribute less overall and gain greater payoffs than those with large intercepts (Figure 4.7). This result illustrates that, as expected for a PD game, the stable degree of cooperation shown in the horizontal transmission model and observed in the experiment is susceptible to selection resulting in evolution towards extinction. For further detail on the simulation model, see appendix 4.4.

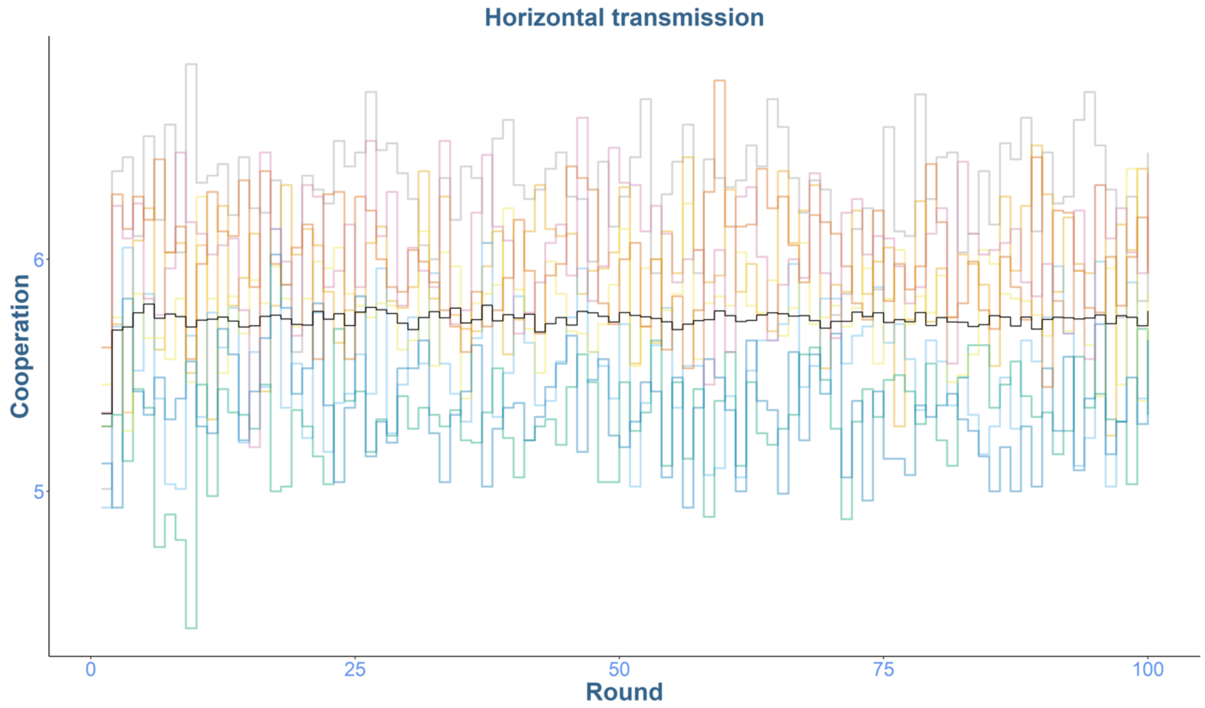


Figure 4.5. *Predicted cooperation across 100 rounds (N = 100) for 100 groups, each consisting of 100 individuals, where, following the experiment, cooperation modifies through horizontal transmission only. Black line shows mean cooperation rates across all groups and coloured lines show cooperation rates for 8 randomly drawn groups.*

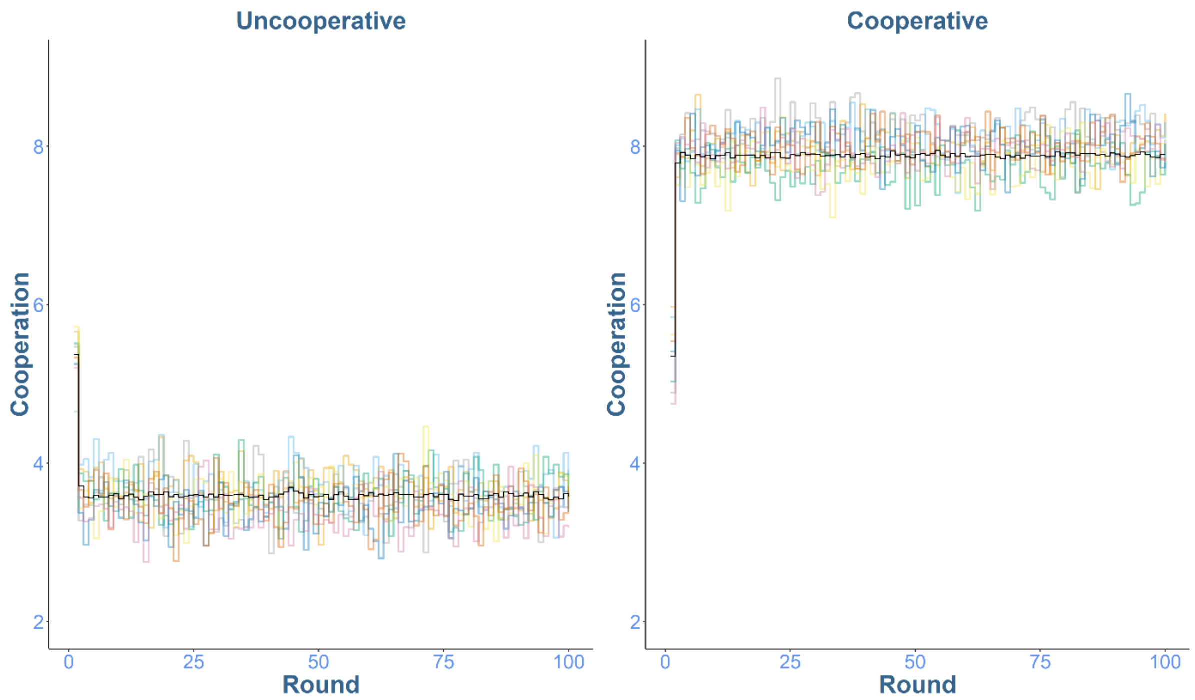


Figure 4.6. *Predicted mean cooperation across 100 rounds (N = 100) for different population compositions as determined by the distribution of intercepts (N = 100). Uncooperative population (left) samples from intercepts below 0 and the cooperative population (right) samples from intercepts above 0.*

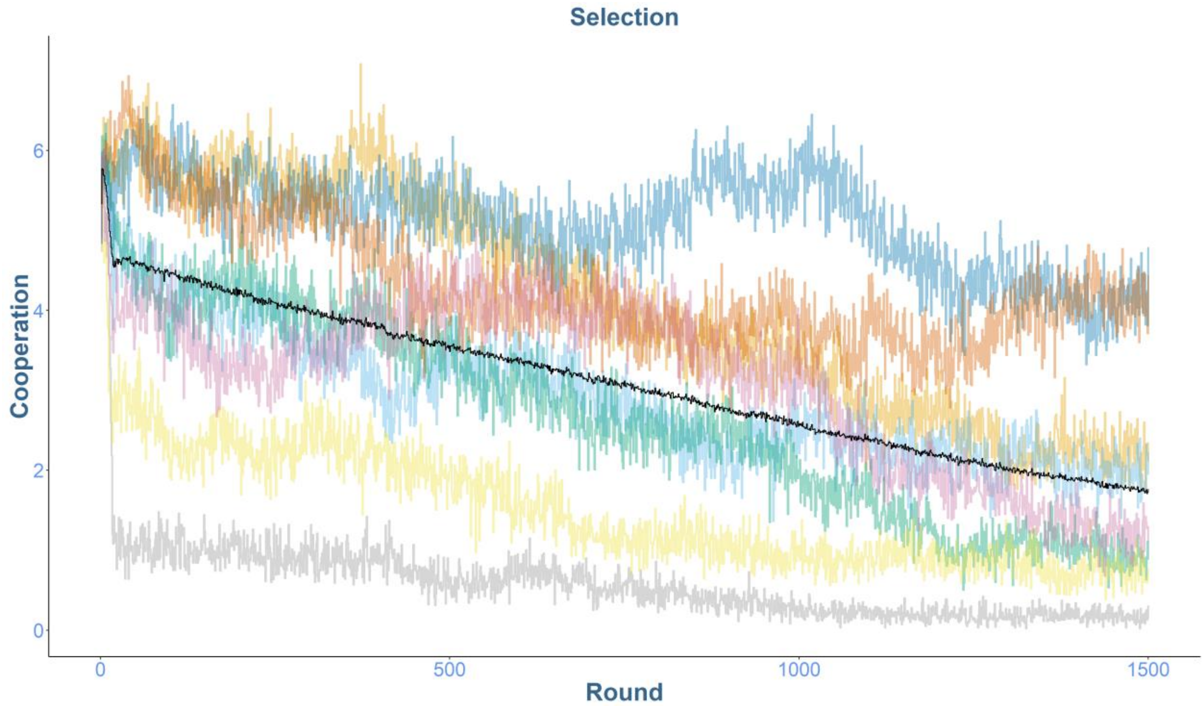


Figure 4.7. *Predicted mean cooperation across 1500 rounds (N = 100) where cooperation modifies by horizontal transmission and intercepts modify through selection and mutation. Black line shows average across 100 groups and coloured lines show cooperation rates for 8 randomly drawn groups.*

4.5. Discussion

This experiment sought to test multiple predictions. Specifically, whether there is evidence for the use of social learning in social dilemmas and if so, which social learning strategy between payoff bias, conformity and prestige bias did participants appear to be following. Further, comparisons were made between Prisoner's dilemma and Snowdrift public goods games. Finally, the statistical estimates of parameters contributing to behaviour in the experiment were fed into a simulation model to predict long-term trends, examining group size and the effects of selection on the propensity for cooperation.

We found evidence for the use of payoff biased learning in social dilemmas, but little support for prestige or conformity. However, the overall impact of the social learning strategies on cooperative behaviour was small. There was little evidence of an interaction between game structure and social learning strategy use. Payoff biased copying has also been found in previous social dilemma experiments where, in each case, social learning and specifically, payoff biased copying eroded cooperation (Burton-Chellew, El Mouden, et al., 2017; Burton-Chellew et al., 2015). These findings add to the growing evidence of payoff biased social learning in a variety of other contexts and species (Barrett et al., 2017; Beheim et al., 2014; Bono et al., 2018; J. R. Kendal et al., 2009; Vale, Flynn, et al., 2017).

In our experiment we found no strong evidence for the use of conformity. Theoretically, conformity may influence patterns of cooperation, but it can often depend on the initial composition of the population (Carpenter, 2004) or other complementary mechanisms such as network reciprocity (Jiang et al., 2015). In social dilemma experiments, conformity can sometimes increase cooperation, though it is outperformed by stronger cues such as reciprocity (Bardsley & Sausgruber, 2005; Romano & Balliet, 2017) is often ignored (Burton-Chellew, El Mouden, et al., 2017), or increases cheating (Diekmann et al., 2015; Gino et al., 2009). Outside of cooperative contexts, frequency dependent information is only used if payoff information is unreliable (McElreath et al., 2008), which may explain our findings. Despite this, a null result in our experiment does not necessarily imply conformity is unimportant for the evolution of cooperation. One of the benefits of strong conformity, often absent from experimental research (Gächter & Thöni, 2005) is the spread of shared cultural norms or values, which in turn, can facilitate cooperation (Molleman, Pen, et al., 2013).

The absence of a strong prestige bias was unexpected. Of the little research available, the effect of prestige or leadership on cooperative behaviour seems overwhelmingly positive (De Cremer & Van Knippenberg, 2002; Gächter et al., 2012; Henrich et al., 2015; Kumru & Vesterlund, 2010; Molleman, Quiñones, et al., 2013). While our study suggests that prestige does not influence cooperation as strongly as other research has suggested, there are several possible reasons for this. Like conformity, it may be that prestige was not used because accurate payoff information was available. By definition, prestige serves as a heuristic to be used when payoff information is ambiguous or unavailable (McElreath & Henrich, 2003), which has been demonstrated in an experimental setting (Brand et al., 2020). It is also important to consider the way prestige was defined in this experiment. A prestigious individual is defined as someone with either high general skill and knowledge and / or with a large following (Henrich & Gil-White, 2001; Jiménez & Mesoudi, 2019). Our operationalization of prestige using a quiz follows other studies that have successfully used this approach (Brand et al., 2020; Brand

& Mesoudi, 2019; Kumru & Vesterlund, 2010). Nonetheless, the possibility remains that our participants did not consider the winner of the quiz to be prestigious in the context of the social dilemma. Moreover, high scoring individuals demonstrated skill in the same domain as the context in which they could be copied (the social dilemma game) rather than a potentially less ‘useful’ general knowledge.

It should be noted that, even for payoff bias, the effects sizes associated with social learning strategies were not particularly large and were all associated with a good deal of uncertainty. This was reflected in the patterns of model comparison which showed only small differences in WAIC scores between competing models, which suggests that each model would make roughly similar out-of-sample predictions. Also, the simulation model indicated that social learning strategies did not cause a significant change in cooperation which, instead, was determined by individual propensities for cooperation (determined by intercepts). Furthermore, when asked, after their participation in the game, whether they used the social information in some way, only 28% of participants (that responded) said yes.

We found lower levels of cooperation in the social information condition than the asocial condition. Although a concern for reputation might suggest that cooperative acts are more common when such behaviour is observable (Kelsey et al., 2018; Yoeli et al., 2013), overwhelmingly, classic economic games which provide breakdowns of group mates’ behaviours, find free riding to be the dominant strategy (R. M. Dawes, 1980; Fehr & Schurtenberger, 2018). In a study which compared playing with and without information about group mates’ behaviour, higher contributions were found in groups where no information was available (Burton-Chellew, El Mouden, et al., 2017; Neugebauer et al., 2009). These, and our, findings suggest that providing social information reduces cooperation. One explanation is that social information is used to update beliefs about how little other group members are contributing (Fischbacher & Gächter, 2010).

Finally, as predicted, we found evidence of higher levels of cooperation in the SD game compared with the PD. Although the effect was small, this result is consistent with existing theoretical work

(Doebeli & Hauert, 2005a; Souza et al., 2009) as well as biological (Gore et al., 2009; R. P. Smith et al., 2019) and experimental evidence (Hilbig et al., 2018; Kümmerli et al., 2007) which predict higher cooperation in SD than PD.

There are several methodological aspects of our study worth addressing. Unlike most other studies which consider SD games using one shot or binary interactions, we allowed contributions on a continuous scale across multiple rounds. The setup of our SD game represented an extremely harsh SD game (e.g., compared to Brown & Vincent, 2008; Sasaki & Okada, 2015) where a failure to meet the public good threshold resulted in a complete loss of all individuals' payoffs for that round. Many formulations (though often binary cases) consider such an outcome to result in no change in individuals' payoffs (Doebeli & Hauert, 2005a). In that sense, our experiment may be more akin to a Chicken game, where mutual defection (or failure to swerve) produces an actively deleterious outcome. Nevertheless, the formulation of our experiment still conforms to the characteristic payoff hierarchy of the SD game (where cooperating against defector(s) is preferable to defecting) which applies to real life contexts. For instance, the failure of a population to reach the investment necessary for functional flood defences or invest sufficiently in predator defence could result in the collapse of that population. Therefore, we maintain that the setup of this experiment is a useful approximation of real-life cooperative dilemmas.

The mean group donation displayed in this experiment was around 6 units (of a possible 10), which showed little decline across rounds. This is unusual for PD social dilemma games, which generally show high initial donations which decline sharply towards the end of the game (Fehr & Schurtenberger, 2018) and average contributions of around 37% (Zelmer, 2003). One possibility is the relatively low number of rounds in our experiment, though previous experiments have shown declines within this timeframe (Burton-Chellew, El Mouden, et al., 2017; Harrell & Simpson, 2016). Alternatively, participants may have been confused about how the game worked (Ferraro & Vossler, 2010). While this is possible, our self-report measure suggested that participants generally believed they understood how the game worked. A more likely explanation for the elevated contribution rate is the multiplication factor of 2 used in this experiment. High multiplication factors have been found to both raise cooperation rates and slow declines across rounds (Gunnthorsdottir et al., 2007; P. van den Berg et al., 2020).

In our experiment, participants could be socially influenced by others taking part in the same iteration of the social dilemma game. This contrasts with other experimental designs which only allow social learning between groups playing different iterations of the social dilemma game (Burton-Chellew, el Mouden, et al., 2017; Romano & Balliet, 2017). The latter approach has benefits, as it allows social learning to be decoupled from other factors such as reciprocity or the possibility of participants attempting to influence their group mates' behaviour through their own behaviour. Nonetheless, we consider that the within-group social influence design holds greater ecological validity in simulating situations where people may be socially influenced by those that are participating in the same social dilemma. The decision not to manipulate what social information was offered to participants also approximated a more realistic scenario, allowing each participant to adopt one or more social learning strategies (R. Kendal et al., 2018; Molleman et al., 2014). Of course, we cannot discount that participants used some other strategy (or combination of strategies) aside from the ones considered here (McElreath et al., 2008).

Future research could address individual differences in social learning strategy use in the context of cooperation (R. Kendal et al., 2018; Rawlings et al., 2017; Vale, Flynn, et al., 2017). A larger sample size than was feasible in our experiment would allow the GLMMs to be extended to include an individual slope for each participant and calculate the proportion of participants who employed a given strategy (Toyokawa et al., 2019). An alternative might be to allow participants to choose what information they viewed (Flynn et al., 2016). Further attention should also be given to prestige as we failed to document a strong effect in contrast to clear predictions from theoretical models (Henrich et al., 2015; Molleman, Quiñones, et al., 2013). To address the possibility that our operationalization of prestige was not relevant to participants in this experiment, it would be useful to consider a different definition of prestige, perhaps one based on popularity (Brand et al., 2020). Experiments could also investigate other game structures than those considered here, such as the stag hunt game (Iyer & Killingback, 2020; Skyrms, 2004). Finally, it would be useful to consider social learning strategies within real-world cooperative scenarios. For example, normative messages are widely used in interventions to reduce household energy use (Allcott, 2011) and cultural group selection has been applied to understand the transmission of lobstering practices in Maine (Waring & Acheson, 2018). Both our study and the literature suggest that payoff bias may affect cooperative behaviour within applied settings. However, given the overwhelming effect of intercept variation in our study, it may also be important to consider factors such as personality and the socio-cultural environment that shapes the development of inclinations to cooperate.

Chapter 5: Social learning does not influence spiteful behaviour: Individuals behave altruistically, but not spitefully, in an online game

5.1. Abstract

While humans are highly cooperative, they can also behave spitefully. Yet, spiteful behaviour remains an understudied aspect of human social behaviour. Real-life spiteful acts can be normatively driven, though no experiments have addressed the direct role of social learning on spiteful behaviour. Here we present an experiment that expands on previous work by allowing participants full flexibility to behave altruistically, neutrally, or spitefully. We also investigate the effect of social information on participant's decisions. We compared two social learning strategies (information source), conformity (most common behaviour) and copy-the-successful (highest earning participant) and varied the content of the social information between altruism, spite or neutral behaviour. We found a marked preference towards altruistic behaviour and weak evidence that this was influenced by altruistic information content, but not the source. Exploratory analysis revealed that participant earnings in a prior phase of the experiment negatively correlated with altruistic behaviour. Our results contrast with previous experiments which find high rates of spite within experimental samples. Theoretical and methodological explanations are discussed.

5.2. Introduction

Humans exhibit levels of cooperation within groups unparalleled in the animal world (Henrich & Muthukrishna, 2021; Kurzban et al., 2015). In experiments, participants frequently display higher levels of cooperation than predicted by game theory (Fehr & Schurtenberger, 2018). Classic explanations include kinship (Hamilton, 1964), reciprocity (Trivers, 1971) and reputation (Nowak & Sigmund, 2005). Recent attention has also been granted to social learning or social norms (D. Smith, 2020). Social behaviour of an actor towards a recipient can be split into four types depending on the payoff consequences (Table 5.1). By this scheme, both altruistic and spiteful behaviour is maladaptive in that they incur a cost to the actor but are distinguished by positive or negative impacts on the recipient, respectively. While West et al. (2007) define, spite as behaviour that is costly to both parties' lifetime fitness, some have suggested the definition should also include behaviour that is neutral to the actor but costly to the recipient (Gadagkar, 1993).

Table 5.1. *Kinds of social behaviour according to fitness consequences to the actor and recipient. Adapted from* (West et al., 2007b).

	Effect on recipient	
Effect on actor	+	_
+	Cooperation	Selfishness
_	Altruism	Spite

While extensive research effort has been directed towards cooperative, altruistic, and selfish behaviour, spiteful behaviour has been historically neglected (Gardner & West, 2004; West et al., 2007b). In his original writings on kinship, Hamilton considered spite in terms of negative relatedness, where spite could be favoured if it was selectively directed at individuals less related to the actor than another random member of the population (Hamilton, 1964). Wilson proposed a similar mechanism, where spite could be favoured if it benefited a related third party (Wilson, 1975). Recent definitions term this *evolutionary spite* (Jensen, 2010). Evolutionary spite is extraordinarily rare, in part because it is difficult to conclusively demonstrate that a given behaviour could not provide direct fitness benefits at a future point (Foster et al., 2001). Others argue that spite may require more sophisticated cognitive mechanisms than is possessed by many animals (Hauser et al., 2009). Theoretically, spite is predicted to be most favourable in environments with intense local competition (Gardner & West, 2004; Jensen, 2010). This is so because when individuals compete within small local groups, rather than between larger social groups, in addition to increased negative relatedness, spite is a more favourable strategy to raise your success relative to a group average (Gardner & West, 2004). This was supported by a theoretical model (Gardner & West, 2004) and an experiment on bacteria, where a spiteful toxin-producing strain was favoured more under local than global competition (Inglis et al., 2011)

An alternative, more relaxed, definition for spite is *functional spite*. This definition includes behaviour that is mutually costly in the short term but could conceivably improve future fitness (Jensen, 2010). This definition makes such behaviour ultimately selfish and includes behaviours such as punishment which is not generally thought of as spiteful (Balliet, Mulder, et al., 2011; Boyd et al., 2003). While spite and punishment are similar in many ways, the motivation to punish is often a change in the target's behaviour or to enforce a social norm (Boyd & Richerson, 1992; Carpenter et al., 2004; Gintis, 2000). For spite, however, the end goal is the harm done to the other individual (Jensen, 2010). In either case, limited examples of spite have been shown in animals and humans.

5.2.1. Spite in animals

Most evidence indicating evolutionary spite in animals comes from social insects. Parasitic wasp eggs laid inside hosts will occasionally forsake their reproductive potential and hatch into sterile soldiers that attack less related individuals (Gardner et al., 2007). Red fire ants who possess the b allele will direct lethal aggression towards queens with the BB allele (Keller & Ross, 1998) and honeybees destroy unfertilised eggs laid by other workers to maintain a reproductive monopoly for the queen (Barron et al., 2001). Though, in the case of fire ants and honeybees, this does not fulfil the strict definition of spite suggested by West et al., (2007) as, besides the energetic costs, these behaviours are not costly to the individuals performing them.

Besides social insects, convincing examples of evolutionary spite have yet to be demonstrated in other animals, because it has not been possible to exclude future direct fitness benefits. Two examples are nest pirating in gulls (Pierotti, 1980) and the harassment of mating pairs in macaque monkeys (Brereton, 1994). Though both behaviours are costly to the actor and recipient in the short term, they may later provide a competitive advantage. Additionally, gull pirating occurs after (rather than causing) the loss of one's own offspring, indicating it is a response to existing poor reproductive success (Pierotti, 1980).

Experimentation may be useful to study spiteful behaviour as this allows fitness consequences to be better controlled. Chimpanzees removed a partner's access to food by collapsing a table if their partner stole their food but not if the experimenters only gave their partner food (Jensen et al., 2007). This suggests that Chimpanzees were engaging in punishment, but not spite. Capuchins, on the other hand, collapsed the table at equal rates irrespective of whether the partner actively stole the food or was simply given it, which is more consistent with spiteful behaviour and inequity aversion than punishment. (Leimgruber et al., 2016). This suggests that the extent of spiteful behaviour may vary between species. In this case, the punishment behaviour (based on uneven resource distribution) exhibited by capuchins is more like patterns observed in humans than chimpanzees (Leimgruber et al., 2016).

5.2.2. Spite in humans

While humans are remarkably cooperative, they can also be remarkably antisocial. For example, in scenarios such as romantic breakups where child custodial agreements end up benefitting nobody or the creation of computer viruses that have little purpose aside from the destruction of data. Humans also often take pleasure in the misfortunes of others in the form of schadenfreude (R. H. Smith et al., 2009). Many violent attacks in response to perceived norm violations can be interpreted as spiteful.

Acid attacks are primarily gang-related, are usually directed at the face (with the intention of humiliating their victims) and are rarely fatal (Lewis et al., 2020). Likewise, suicide bombings are often motivated by desires for revenge (Gambetta, 2005). In both cases, the goal is the harm to the victim, rather than affecting change in the target's behaviour, which is consistent with functional spite (Jensen, 2010). There are limited ethnographic accounts of spite, for example Ceylon village residents would often report neighbouring houses to a commission to be torn down without any clear incentive (Leach, 1961) but little work has focused on it explicitly. It should also be remembered that these cases are not necessarily evidence of evolutionary spite (West et al., 2007b), as the lifetime fitness consequences of such behaviour cannot be accurately assessed.

Laboratory studies may offer further insight into the case of humans. While not usually thought of as spiteful, rejecting offers in the ultimatum game could be interpreted as such: here a player can choose to reject a proportion of a fixed fund offered by another player, causing both players to lose out (Güth & Kocher, 2014). A key motivation is anger at the inequity, which has been shown to mediate rejections within the ultimatum game (Srivastava et al., 2009) and motivate the use of punishment (Pedersen et al., 2018). In an experiment where participants witnessed an unfair exchange between individuals A and B, they readily punished B at a personal cost, offering reasons like "wanting B to lose" (Fehr et al., 2008). This also occurred when participants received randomly generated scores where participants paid points to reduce the score of high earners and compensate low earners, expressing anger at those who had more points than they did (C. T. Dawes et al., 2007). These examples may align closer with functional spite than evolutionary spite, as harm is done here for its own end rather than a desire to motivate future behaviour.

Spiteful behaviour has been investigated in a variety of other ways, though it is not always costly to the actor. One experiment allowed participants to make their partner's search task easier or harder by manually adding or removing characters. Around 37% of participants chose to harm their partner to at least some degree (Sadrieh & Schröder, 2017). In a mock auction where the item was sold to the highest bidder at the second-highest price, participants could behave spitefully by driving the purchase price up without risking winning the item. In this case, participants were split bimodally, 70% were never spiteful but the remainder were maximally spiteful (Kimbrough & Reiss, 2012).

Further evidence for spite comes from a collection of experiments specifically designed to assess it. In the first experiment of this kind, participants could pay a portion of their earnings to reduce their partner's earnings, which around 62.5% of participants did to some extent (Zizzo & Oswald, 2001). In the Joy of Destruction game, up to 40% of participants chose to reduce their partner's earnings when there was a 1/3 chance that both participants would lose their earnings anyway (Abbink & Sadrieh, 2009), but this dropped to 25% when reducing points was costly to the actor and to 10% when decisions could not be hidden behind the random removal of points (Abbink & Herrmann, 2011). Spite was less common (38% versus 18%) when participants made decisions in the presence of eyes (Baillon et al., 2013) or after receiving a hug (22 % versus 0%), but spite did not correlate with personality (Blackwell & Diamond, 2017). Both spite and altruism were less common when participants were offered both options simultaneously rather than separately and many participants behaved inconsistently (behaving altruistic and spiteful) between games (Zhang & Ortmann, 2016). Though, when presented one at a time, altruistic behaviour did positively correlate with spiteful behaviour and additionally with sensitivity to social pressure (Zizzo & Fleming, 2011). Outside of the laboratory, in a field experiment with farmers in Namibia, around 40% (versus 23%) were spiteful in regions of poor crop yield supporting predictions that spite is a response to conflict (Prediger et al., 2014).

Overall, these experiments suggest that humans do exhibit spite in several contexts, which is surprising given the degree that we also cooperate (Henrich & Muthukrishna, 2021). One explanation is that spite evolved as a by-product of costly punishment as it relies on the same cognitive mechanisms (Hauser et al., 2009). This is plausible, as animals generally do not demonstrate spite or punishment bar a few specific examples (Arseneau-Robar et al., 2016; Bshary & Grutter, 2002). Additionally, spite is more common alongside intense local conflict or when there are large imbalances between individuals' earnings (Gardner et al., 2007; Gardner & West, 2004; Prediger et al., 2014).

However, there are several methodological limitations of experiments that implement the Joy of Destruction game. In some studies, spiteful behaviour is not costly to the actor (Abbink & Sadrieh, 2009; Blackwell & Diamond, 2017; Zhang & Ortmann, 2016). Spite being costless is inconsistent with the definition of evolutionary spite (West et al., 2007b) and may also inflate experimental estimates of its prevalence. Consistent with this, participants were more cooperative in a hypothetical prisoner's dilemma game than in an incentivised one (Lönnqvist et al., 2011). Furthermore, in all studies except Zhang and Ortmann (2016), participants were restricted to behaving spitefully or doing nothing (Abbink & Herrmann, 2011; Abbink & Sadrieh, 2009; Baillon et al., 2013; Blackwell & Diamond, 2017; Prediger et al., 2014; Zizzo & Oswald, 2001) or were offered the choice to be altruistic or spiteful one at a time (Zizzo & Fleming, 2011). Spite may be common in these cases because it is novel but may conflict with some participants' preferences for compensation rather than punishment (Pedersen et al., 2018). To date, no study considering costly spite has offered participants the full range of possible behaviour from spite to altruism. Addressing this limitation is a key aim of our study.

5.2.3. Social Learning Strategies

The proximate factors that may explain spiteful behaviour are also not well understood. One potential explanation is that spiteful behaviour may spread through social learning. Humans are adept social learners (Boyd et al., 2011; Legare, 2017) which allows the quick acquisition of difficult to obtain information and the formation of cumulative cultural traits (Mesoudi & Thornton, 2018). Unbiased copying is unlikely to be favoured by natural selection as it risks obtaining outdated or poor information (Boyd & Richerson, 1995). For this reason, it has been proposed that social learning strategies would be adaptive, influencing when what and from whom individuals socially learn (R. Kendal et al., 2018; Laland, 2004; Morgan et al., 2012). Examples include a bias towards types of information (for example, social content (Tehrani et al., 2015)), or model-based biases such as conformity (Henrich & Boyd, 1998; Whiten, 2019a) success (McElreath & Henrich, 2003; Sarin & Dukas, 2009) or prestige (Henrich & Gil-White, 2001; Jiménez & Mesoudi, 2019). The effects of social learning strategies on spiteful behaviour have yet to be examined (Baillon et al., 2013; Zizzo & Fleming, 2011), but there are reasons to speculate that spiteful behaviour may be copied by individuals.

Firstly, participants have been shown to increase their cooperation in response to either an individual of high status (Kumru & Vesterlund, 2010) or individuals from another group who are behaving cooperatively (Romano & Balliet, 2017). Within a field context, normative prompts are often used to encourage reductions in energy use (Allcott, 2011). Conversely, there is evidence that individuals use a payoff bias in social dilemmas, which seems to be favoured over other social learning strategies (Burton-Chellew, El Mouden, et al., 2017; Burton-chellew & Amico, 2021; Molleman et al., 2014; Watson et al., 2021). The results of study 1 in this thesis (chapter 4) indicated that the changes in participant's cooperation was most consistent with a payoff bias over conformity or a prestige bias. Given that social learning is used to inform cooperative decision making, it is plausible that it may also influence spiteful behaviour.

Social learning also overcomes a major barrier to the evolution of spiteful behaviour; if one puts aside kinship, spiteful behaviour is maladaptive, as individuals who engage in spiteful behaviour are invariably worse off than those that do not. It has been shown that social learning strategies can sometimes result in the spread of maladaptive behaviour, particularly when the adaptive value of the adopted behaviour is not directly assessed by the copier (Giraldeau et al., 2002; Mesoudi, 2009). For example, a bias to copy generally successful individuals can include behaviours that are irrelevant to the copied individual's success (Henrich & Gil-White, 2001). Likewise, a bias to copy the most frequent behaviour in a population can result in the spread of sub-optimal behaviours (Kandler & Laland, 2009).

Finally, there is evidence that punishment is copied by individuals. Experimental evidence has shown that individuals change their preference towards punishment if they are provided feedback that their group mates favoured punishment (FeldmanHall et al., 2018) or if information suggests punishment and cooperation is the normative behaviour (Li et al., 2021).

Within competitive sports, studies have found cases of the social transmission of cheating which, similar to spiteful behaviour, is commonly considered antisocial behaviour. Following baseball player Jose Canseco's confirmed use of performance-enhancing drugs, later analysis showed that this correlated with marked improvements in his teammate's performances which is suggestive that drug use was socially transmitted within the team (Gould & Kaplan, 2011). In football, antisocial behaviour in the form of intentional fouling or aggressive play was shown to be predicted by associating with peers or coaches who endorsed it (Kabiri et al., 2020; Malete et al., 2013). These cases are also interesting because they occurred in environments of extreme competition, but they generally point to the possibility that spite can be influenced by other individuals.

Here we consider two well studied social learning strategies, conformity (Henrich & Boyd, 1998; Whiten, 2019a) and copy-the-successful (McElreath & Henrich, 2003; Sarin & Dukas, 2009). Theoretically, conformity is adaptive if beneficial traits are present in the population at the greatest frequency (Henrich & Boyd, 1998). Individuals using conformity are disproportionately likely to adopt the most common trait (Henrich & Boyd, 1998; Whiten, 2019a). However, populations can fixate on suboptimal traits if conformity is too strong (Kandler & Laland, 2009; Whitehead & Richerson, 2009). The use of conformity has been found in adult (Efferson et al., 2008; Muthukrishna et al., 2016) and non-adult (Morgan et al., 2015) participants though other social learning strategies such as payoff bias are more likely to be employed (Barrett et al., 2017; McElreath et al., 2008). Provided it is possible to assess the relative payoffs of behaviours or individuals (McElreath & Henrich, 2003), copying traits with a probability proportional to their payoffs can outcompete other strategies (Schlag, 1999). A copy-the-successful bias has been shown in a variety of contexts, including stickleback feeding preferences (J. R. Kendal et al., 2009), the transmission of food extraction methods in capuchins (Barrett et al., 2017) and managers' formation choices in premier league football (Mesoudi, 2020).

5.2.4. Research questions

Our study seeks to evaluate the extent that spiteful behaviour occurs when it is costly to the participant and when participants are also offered the choice to be altruistic. We will also test whether social learning strategies affect the incidence of spiteful behaviour, focusing on two well-studied social learning strategies: conformity and copy-the-successful (Henrich & Boyd, 1998; R. Kendal et al., 2018; McElreath & Henrich, 2003). We address the following research questions (RQ):

RQ1. To what extent is spiteful behaviour exhibited in our experiment? - From experimental literature, we predict between 10%-40% of participants will behave spitefully.

RQ2. Does social information enabling the use of conformity or copy-the-successful strategies affect altruistic or spiteful behaviour? – There is stronger evidence for the effect of payoff-biased learning than for conformity on cooperative behaviour, so we predicted that copy-the-successful information will exert a stronger influence than conformity information on participant's behaviour.

5.3. Methods

5.3.1. Design

After playing a game where participants earned points (part 1), participants were given either specific forms of social information, specified below (Table 5.2), or assigned to an asocial control group, and were given the opportunity to donate (altruism) or withdraw (spite) points from an anonymous partner at a cost to themselves (part 2).

We ran 6 social information conditions in a between-participants 3x2 factor design (Table 5.2). Unknown to the participants, the information participants were presented with was fictitious. Factor 1 was the source of social information (the majority of previous participants or the most successful previous participant), while Factor 2 specified the behaviour of the source towards their partner (spite, altruism or neutral). In addition, we included an asocial control condition, where participants did not receive any social information.

Table 5.2. Social information conditions. Social information presented to participants varied by the source of the information among previous participants (Factor 1) and the source's behaviour towards the partner player (Factor 2). All social information was fictitious but perceived to be real by the participants. Numbers in brackets show the sample size in each condition. The control condition not displayed in the table contained 54 participants.

Information source (Factor 1)	Source's behaviour (Factor 2)		
Successful	Reduced points of partner (47)	Did not change points of partner (57)	Increased points of partner (41)
Majority	Reduced points of partner (53)	Did not change points of partner (43)	Increase points of partner (51)

5.3.2. Materials and procedure

The experiment was conducted using the experimental platform Dallinger (Dallinger, 2020) and participants were recruited on Amazon's Mechanical Turk (MTurk). Once participants joined the experiment, a screen indicated they were awaiting a second participant. After a short delay, the experiment began. Throughout, participants were deceived into thinking an anonymous second participant was simultaneously taking part in the experiment. To enhance believability, randomised timed delays were used throughout the experiment to suggest they had to wait for the other participant to catch up. The experiment was split into 2 parts and participants completed both parts.

In Part 1 (see appendix 5.1), participants knowingly played a 5-round game with a bot. The purpose of this was for participants to accumulate points to be used in part 2. It was important for participants to feel they had earned their points to avoid any "house-money" effects, where participants are more reckless with points or money, they do not feel is theirs (Abbink & Sadrieh, 2009; Harrison, 2007). Participants were told that the points they had obtained by the end of the experiment would be converted to a bonus payment (participants were not told how much the points were worth).

In each round of part 1, participants were given 10 points and could send any amount of this to the bot. The bot then returned a sum of points which varied based on how much the participant sent. Their round score was determined by the points received from the bot plus the points they kept for themselves. The point values returned from the bot were randomly varied within a large range to prevent participants from easily working out the pattern.

In part 2 (see appendix 5.2), participants were told that either they or the other participant would be assigned a "decider" role and could change their other participant's score. It was made clear that this was a one-shot decision, and the recipient would have no opportunity to respond to the decider's choice. Unknown to the participant, they were always assigned the "decider" role while their partner was, in reality, a bot. In each of the social conditions, participants received fake information about a different previous participant's score-change decisions. Depending on the experimental condition (Factor 2), social information stated that previous participants either: "did not change their partner's score" (N), "increased their partner's score" (I) or "decreased their partner's score" (D). Depending on the experimental condition (Factor 1), the source of the information was stated to be either "the majority of previous participants" (conformity) or "the highest scoring participant in previous games" (copy-the-successful).

Participants could change their partner's score in increments of 3 (maximum 30) for every 1 point they paid using a slider, which updated to show their new score and how much their partner's score would increase or decrease. This ratio was chosen based on previous studies employing costly punishment (Fischbacher & Fehr, 2004; Rand & Nowak, 2011). This represented a monetary cost for participants, as their points at the end of the experiment were converted into a bonus payment.

Following part 2, we collected free-text responses to gain insight into participant's decision making. To act as a comprehension check, participants were also asked to specify if they had chosen to increase, decrease or not change their partner's earnings. Participants were then debriefed, and the deception employed in the experiment was explained. They were reminded of their right to withdraw at this point (though, very few did). Finally, demographic information was collected, and participants were asked to rate their level of understanding of the game.

5.3.3. Participants

Participants were recruited from MTurk in blocks of 75 and completed the experiment online using their own devices. Each participant was assigned to a condition using a random number generated in JavaScript. Participants who had not completed the experiment or those that had requested their data be removed were excluded, leaving 346 participants. Because conditions were assigned randomly, there was some imbalance between conditions (Table 5.2). Due to a software error, 2 participants had 2 responses associated with their ID. In these cases, the first response (as determined by time created) was kept and the other observations were discarded.

Of those who provided demographic information, the median age was 32 years (IQR = 9) with 197 males, 75 females and 2 non-binary individuals. 253 participants identified as White, 28 Asian, 34 Black African or Caribbean, 12 Latin American, 6 mixed and 3 withheld this information. All participants earned a minimum of $0.35 dollars for completing the experiment with a further $0.60 dollars earnable as a bonus. Participants earned $0.65 on average and the experiment took around 5 minutes to complete.

5.3.4. Data analysis

Analyses were conducted in R studio version 4.1.0 (R Core Team, 2021). Data tidying and plotting were done using *ggplot2* and *tidyverse* and statistical analysis using *brms* (Bürkner, 2017). We termed the outcome variable ‘social behaviour’, which was an integer between -10 and 10 representing a participant’s decision as the decider. 10 indicated they had increased their partners score by the maximum amount and -10 that they had decreased their partners score by the maximum amount. This variable was treated as continuous. MCMC (Markov Chain Monte Carlo) was used to construct the posterior distributions and the models used 4 chains of 1000 warmup samples and 3000 real samples for inference. Model diagnostics indicated excellent chain convergence (all Rhat values = 1) and all effective sample sizes were above 3000. All models used weakly regularising normally distributed priors (see below). Data and code used for plotting and analysis are available here (https://osf.io/ekmuj/?view_only=b65423eccd58496ea05e1bd6b1b0b764).

To address RQ1, we fit an *intercept model* to the data. This model contained only a single intercept parameter and therefore returned a single posterior distribution corresponding to the average social behaviour exhibited by participants across experimental conditions. To address RQ2, we fit two models: a *control model* which included a binary variable to evaluate the effect of participants' exposure to social information in comparison to the asocial control ("Control"); then we fit a *condition model* to evaluate the social information effects by including a varying intercept which partially pooled information across conditions.

Finally, we conducted an exploratory analysis to evaluate the extent to which participant's score in part 1 affected their social behaviour (altruism, spite or neutral behaviour). While RQ1 and RQ2 were specified at the research design phase, this research question was specified after examining the data to test RQ1 and RQ2. This analysis used a *score model* that included participant's score in part 1 as a linear fixed effect. Note that this model also included the varying condition intercept. All models are captured by the following structure:

$$Social\ behaviour \sim Normal(\mu, \sigma)$$

$$\mu = \alpha_{condition} + \beta_{score} * Part\ 1\ Score$$

$$\alpha_{condition} \sim Normal(\bar{\alpha}, \sigma_{\alpha})$$

$$\bar{\alpha} \sim Normal(0,4)$$

$$\beta \sim Normal(0,2)$$

$$\sigma \sim Exponential(1)$$

$$\sigma_{\alpha} \sim Exponential(1)$$

To assess the relative out-of-sample predictive ability of competing models, widely applicable information criteria (WAIC) was computed from the log likelihoods of each model and includes a parameter penalty to balance against the risk of overfitting. Lower WAIC values indicate lower out-of-sample deviance and correspondingly higher predictive value. Posterior predictions are generated to visualise the predicted effects. See McElreath (2020) and Kruschke (2015) for further discussions on this approach and the implementation of Bayesian models.

5.4. Results

5.4.1. To what extent is spiteful behaviour exhibited in our experiment? (RQ1)

Very little: the *intercept model* showed a distribution of positive (altruistic) intercept values best explained the observed social behaviour towards a partner across conditions (Figure 5.1a; mean: 2.77; 95% PI: 2.42 - 3.13 and SD: 3.38; 95% PI: 3.13 - 3.64). In addition, the overall descriptive frequency of altruistic behaviour (66.47%) was far higher than neutral (25.14%) or spiteful behaviour (8.38%) across conditions (Figure 5.1b).

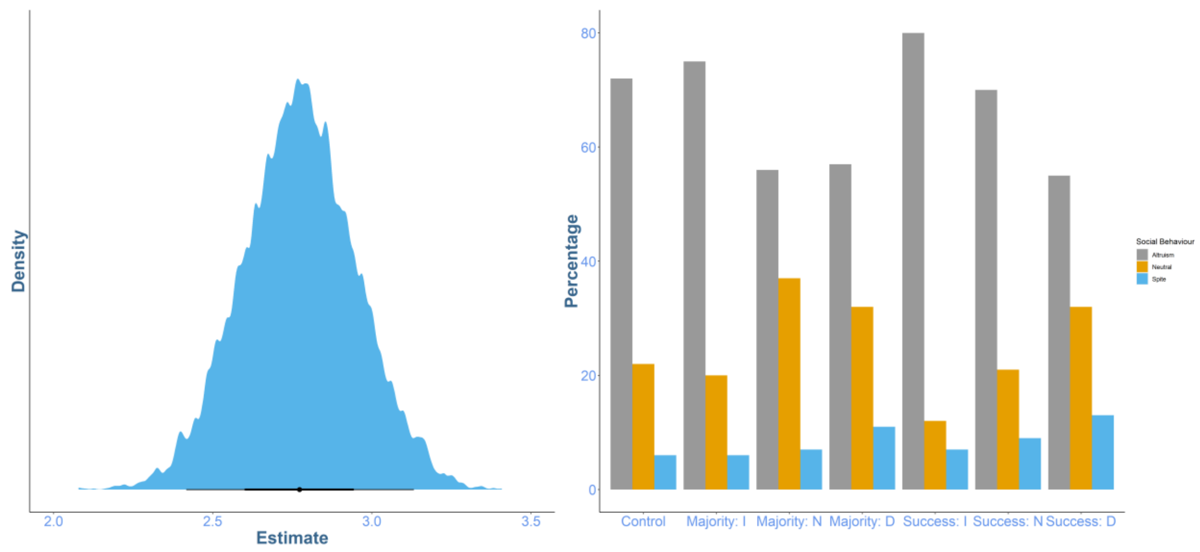


Figure 5.1*. *Left: *Posterior distribution from the "Intercept" model predicting social behaviour exhibited by participants across experimental conditions. Point indicates the mean and lines indicate the 68% and 95% prediction intervals of the sample. Note, positive numbers indicate altruistic behaviour.* ***Right:*** *Percentage of participants within experimental conditions opting for altruistic (grey), neutral (yellow) and spiteful (blue) behaviour. Letter following the colon in the x axis label indicates the type of information seen by participants: I = Increase points, N = No change, D = Decrease points.*

5.4.2. Does social information enabling the use of conformity or copy-the-successful strategies affect altruistic or spiteful behaviour? (RQ2)

No: social information enabling the use of conformity or copy-the-successful did not affect altruistic or spiteful behaviour. The *control model* predicted similar social behaviour in the control and social conditions (Control; M = 2.99, 95% PI = 2.14; 3.85, Social; M = 2.74, 95% PI = 2.36; 3.12) and was not favoured by WAIC compared to the *intercept model* (WAIC scores: Intercept = 1827.7, SE = 27.1, weight = 0.67; Control = 1829.1, SE = 27.0, weight = 0.33). This indicates that participants generally behaved equivalently between the asocial condition and the social conditions. We also note high uncertainty (see SE in WAIC scores) in the estimation of the out-of-sample predictive value of each model.

The *condition model* showed that while the source of the social information (information enabling conformity or copy-the-successful) had no discernible effect on participant's behaviour, there was weak evidence for an independent effect that exposure to social information indicating previous participants had behaved altruistically resulted in more altruistic behaviour than exposure to spiteful or neutral information. The condition model was only marginally favoured by WAIC, indicating both models would perform similarly out-of-sample (WAIC scores: Intercept = 1827.7, SE = 27.1, weight = 0.46; Condition = 1827.4, SE = 28.1, weight = 0.54). Predictions for the intercept parameters from the *condition model* are shown below in Figure 5.2.

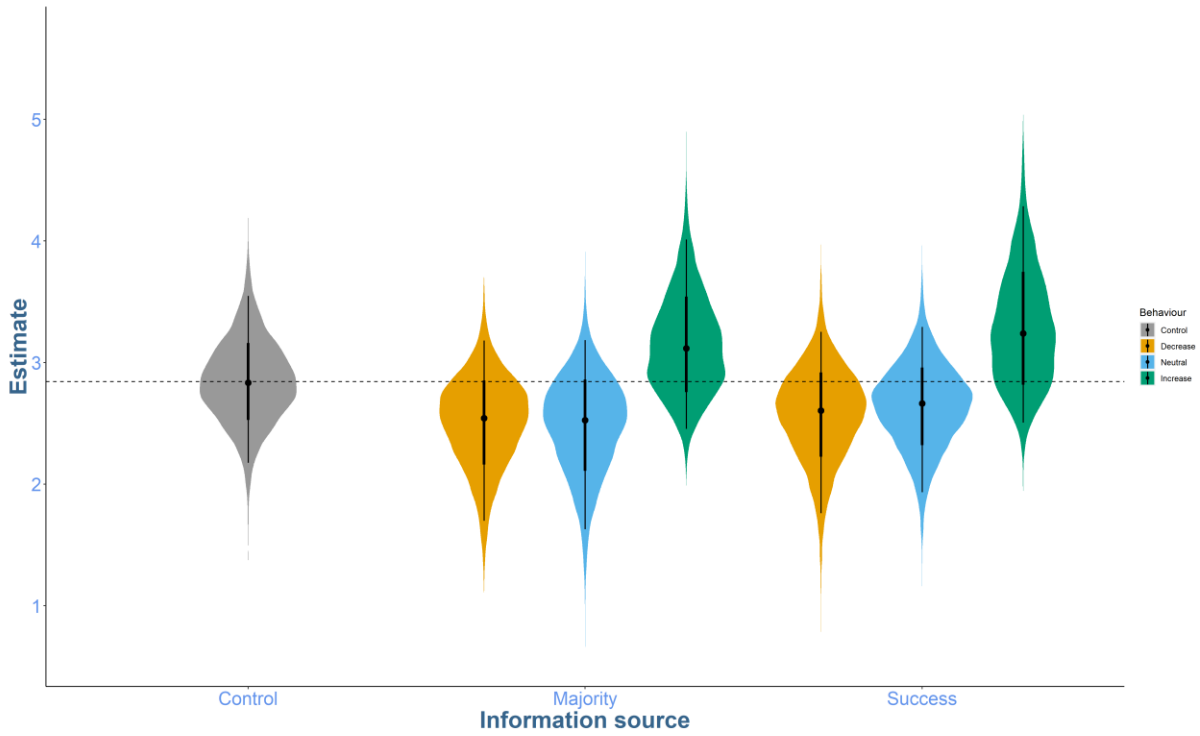


Figure 5.2. *10000 samples of intercept parameters across experimental conditions drawn from the posterior distribution of the condition model predicting social behaviour. Points show the mean of the sampled distribution, and the surrounding lines display the 68% and 95% prediction intervals. Colours indicate the behaviour of individuals in the social information source in terms of whether they decreased (yellow), increased (green) or did not change (yellow) their partner's score. The dashed line indicates the control condition mean, displayed for comparison.*

5.4.3. Exploratory analysis of the influence of participant's earnings on social behaviour

One further model evaluated the extent that participant's score in part 1 (their success) influenced their decisions as the decider. Model comparison supported the inclusion of score into the model as models that included score accounted for 93% of the WAIC weight from the three models (Table 5.3). This suggests that variation in social behaviour was far more strongly associated with their part 1 score than with the social information condition to which they were assigned. Predictions from the *score model* (Figure 5.3) indicated that participants who earned more in part 1 tended to be less generous in part 2 than those who earned less in part 1. The relatively larger uncertainty associated with high participant scores is likely a result of the low frequency of participants that acquired high scores.

Table 5.3. WAIC values and model weights for the three models fit to the data. Standard error difference provides the standard error of the difference between each model and the model with the lowest WAIC score while standard error indicates the standard error of the associated WAIC score.

Model	*WAIC*	*SE*	*Standard error difference*	*Weight*
Score	1821.1	29.5	0	0.93
Condition	1827.4	28.1	2.9	0.04
Intercept	1827.7	27.1	3.4	0.03

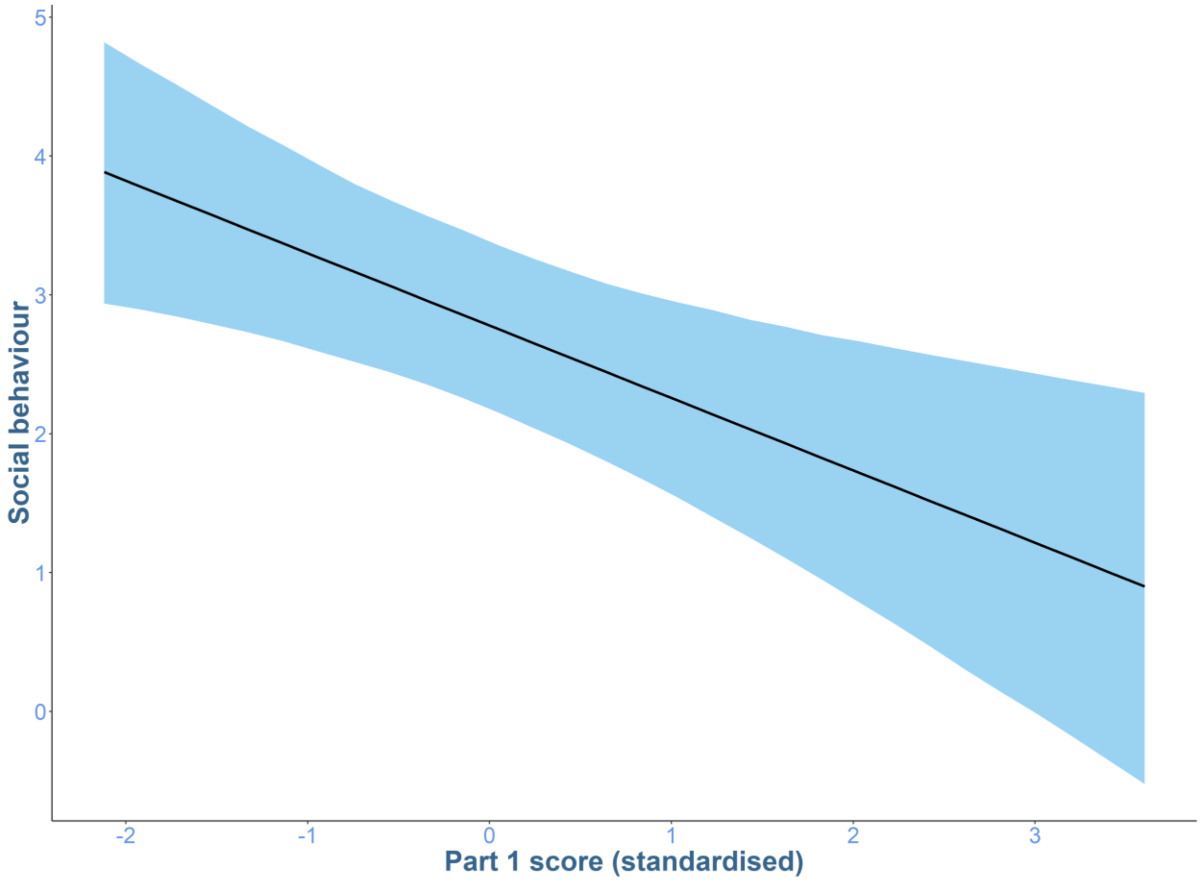


Figure 5.3. *Social behaviour averaged across experimental conditions predicted by a participant's score in part 1. A high score indicated a higher bonus payment at the end of the experiment. The line shows the mean of the posterior predictions, and the shaded region represents the 95% PI.*

5.4.4. Participant's understanding of the experiment

Participants self-report ratings indicated a generally good understanding of how the experiment worked (rated out of 10: median = 8, IQR = 3) which suggested participants did not feel confused during the experiment. Participants were also asked to report what they had done as the decider. Of the 224 participants that provided a response: 99/112 (88.39%) participants correctly reported they had increased their partners score, 61/86 (70.93%) that correctly self-reported they had done nothing but only 10/26 (38.46%) correctly reported they had decreased their partners' score (many believing they had increased it). Though this may be a symptom of the fact that only a small number of participants did opt to reduce their partners' score, this indicates some confusion about how the decider role worked, specifically with respect to spiteful decisions. Therefore, intentional rates of spite may be lower than indicated by this study.

5.5. Discussion

Our experiment investigated two research questions: the prevalence of costly spiteful behaviour when participants could also behave altruistically (RQ1) and the extent that participants' behaviour was influenced by social information enabling the use of conformity or copy-the-successful (RQ2). Overall, we found extremely low rates of costly spite, as participants preferred to behave altruistically and weak evidence that altruistic social information increased altruistic behaviour which did not vary with the information source. In an exploratory post hoc analysis, we found a negative relationship between a participant's score in part 1 (their success) and their altruism.

The extremely low rates of spite in our experiment (8% across experimental conditions), while not 0 as game theory would predict, is far lower than levels found in many previous experiments (Abbink & Sadrieh, 2009; Baillon et al., 2013; Prediger et al., 2014; Zizzo & Fleming, 2011). Our results were more like the hidden condition reported in Abbink & Herrmann (2011) where spite could be masked behind random destruction and Zhang & Ortmann (2016) who found preferences for altruism over spite. Therefore, our results were unexpected but are somewhat consistent with the lower bounds of spite found in some experiments.

There are several possibilities that may explain our results: mostly, these relate to differences between the design of our experiment and others that have investigated spite. Our primary goal was to remove any motivation for participants to make any changes to their partner's score aside from their desire to benefit or harm another individual. This was to represent the theoretical definition of spite as closely as possible, where harm is done for its own end at a cost to the actor (Jensen, 2010). In many previous studies, spite was not costly to the actor (Abbink & Sadrieh, 2009; Blackwell & Diamond, 2017; Kimbrough & Reiss, 2012; Zhang & Ortmann, 2016). This is not consistent with the definition of evolutionary spite (West et al., 2007b), though it would still qualify as functional spite (Gadagkar, 1993; Jensen, 2010) and it may not necessarily reflect how participants would behave if their behaviour was incentivised (Lönnqvist et al., 2011). Furthermore, some studies do not permit participants free choices between altruism, spite or neutral behaviour (Abbink & Herrmann, 2011; Baillon et al., 2013; Blackwell & Diamond, 2017; Prediger et al., 2014), or they investigate spite and altruism in separate games (Zizzo & Fleming, 2011). This is relevant, as some participants prefer to compensate low earners rather than punish high earners (Kimbrough & Reiss, 2012; Pedersen et al., 2018; Zhang & Ortmann, 2016). Taken together, our results suggest that previous studies may have overestimated the prevalence of spite, as we did not replicate the high rates of spite found in previous studies.

Alternatively, our study may have been too restrictive, or lacked sufficient ecological validity, to produce rates of spite comparable to other studies. A key ingredient for spiteful behaviour is an environment of local competition (Gardner & West, 2004), which characterises the examples of spiteful behaviour observed in animals (Gardner et al., 2007; Keller & Ross, 1998). This study took place online and intentionally avoided any cues or language suggesting competition, including not disclosing the score of the participant's "partner". In previous studies, the experimental paradigm was inherently competitive (Kimbrough & Reiss, 2012) or was conducted in person which (inadvertently or otherwise) may have induced a more competitive mindset in participants. Furthermore, anger is a key motivator behind punishment behaviour (Fehr et al., 2008; Pedersen et al., 2013) and other sanctioning behaviour such as rejections in the ultimatum game (Srivastava et al., 2009). Anger was unlikely to be a large part of this study, given that a participant's partner had nothing to do with their score. The lack of these two factors could explain the lack of spite observed in this study. If so, this would suggest that spite does not occur for no reason, as some other studies have suggested (Abbink & Sadrieh, 2009). At least some motivation such as anger or resource inequality may be required for individuals to engage in spite.

Our estimation of spite may be inflated by participant confusion (Ferraro & Vossler, 2010). Of the 26 individuals who exhibited spite in our experiment and responded to the free text questions, only 10 correctly reported they had been spiteful. Many erroneously believed they had been altruistic. As such, true intentional rates of spite in our experiment may be lower still than 8%, which places our results further from the rates of spite observed in other research. Having said this, participants self-reported a good level of understanding of how the experiment worked (rated out of 10: median = 8, IQR = 3) and participants who were altruistic or neutral more accurately self-reported their behaviour choice. It is not clear from the data collected in this study why spiteful behaviour specifically induced more confusion.

The high level of altruism expressed in our sample requires explanation. While it is possible that MTurk participants display high rates of altruism, perhaps because they feel part of an ingroup of fellow MTurk workers, a previous review indicated that economic game results from MTurk samples are comparable to those conducted in person, which typically find low rates of altruism (Rand, 2012). Therefore, these results are probably not explainable through the sampling method. The altruism in our study is more consistent with other one-shot games than the Joy of destruction or money burning games. Offerings in the dictator game for example are on average around 28% (Engel, 2011) which echoes the average social behaviour observed in this experiment (8.39 – 27.97%). Offerings in the trust and ultimatum games are higher (around 50% Güth & Kocher, 2013; N. D. Johnson & Mislin, 2011) but these games give receivers the chance to respond, so larger offerings are expected.

Altruism was negatively related to participants' scores: participants who scored higher in part one were less altruistic. Previous studies have found mixed results regarding income heterogeneity. Some have shown negative relationships (Buckley & Croson, 2006; Erkal et al., 2011), while others find no relationship (Anderson et al., 2008). One potential explanation is entitlement, as participants earned their sum of money this may make them feel like they deserve to keep it or that having earned their points granted them some form of status (Erkal et al., 2011). Previous studies which have considered the wealth (point total) of players found individuals were more spiteful to wealthier players (C. T. Dawes et al., 2007; Zizzo & Fleming, 2011) but no relationship between being wealthy and being spiteful (Zizzo & Oswald, 2001). Spiteful behaviour exhibited towards wealthier individuals appears to be motivated by the negative emotions associated with income inequity (Dawes et al., 2007; but see Marczyk, 2017).

With respect to social learning, there was very little evidence in this study that the source of social information affected participants' decision making, but weak evidence that exposure altruistic social information enhanced altruistic behaviour. Neither frequency information (conformity) (Henrich & Boyd, 1998) nor success weighted information (copy-the-successful) (McElreath & Henrich, 2003) appeared to be influential in this experiment. Furthermore, of the 205 responses in the self-report questionnaire, only 15 participants indicated that the social information had any effect on their behaviour. This was also unexpected given that previous studies have shown punishment norms are transmitted between group mates (FeldmanHall et al., 2018; Li et al., 2021) and that social learning can result in the spread of maladaptive traits particularly when heuristic strategies are relied upon (Mesoudi, 2009; Whitehead & Richerson, 2009). The lack of an effect associated with social learning in this case, suggests that the social information was not sufficient to overcome participants intrinsic desires to behave altruistically.

There are several additional factors worth highlighting here. Theoretically, social learning is most useful when acquiring traits asocially is costly or difficult (R. Kendal et al., 2018). As this was not the case in this study, this may explain why participants did not use any kind of social learning. Though the effect was extremely small, of the differences seen between experimental conditions, there was an equivalent increase in altruism in response to an altruistic majority (Majority: I) and a successful altruistic individual (Successful: I). Likewise, altruism decreased slightly in all cases (irrespective of information source) where the social information suggested previous participants had been spiteful or had not changed their partner's score compared to the control condition. That is, participants responded equally to conformity and copy-the-successful. This suggests that participants were somewhat sensitive to the content of the social information but not the source. They may therefore have employed a content over a context (model-based) bias (Laland, 2004). This might also be explained by the fact that participants had complete information about the payoff consequences of different behaviours, as theory suggests that heuristic model-based strategies (see chapter 2) are only used when payoff information is unavailable (McElreath & Henrich, 2003).

5.6. Conclusion

In sum, this study failed to replicate the results of previous studies regarding spiteful behaviour, finding far less use of spite and strong preferences towards altruism. Social information had very little effect on participants' behaviour: there was no evidence that participants' decisions were influenced by majority or success weighted social information. Our experiment suggests that spiteful behaviour may not occur for no reason and that previous experiments may have overestimated its prevalence. It may require anger induced from income inequality or a competitive scenario to manifest. Alternatively, participants may generally prefer to benefit their partner when offered the choice. Further studies could continue to refine the spite paradigm by ensuring to provide both options (spite, altruism) to participants for a cost (to ensure consistency with theoretical definitions) but also systematically consider important cues such as competition. This study is the first exploration of the potential role of social learning to explain spiteful behaviour, so further studies may yet uncover interesting correlates with this understudied social behaviour.

Chapter 6: Social learning and sustainability in a student population: A mixed methods investigation

6.1. Abstract

Sustainability remains an important part of 21st century life. Although positive environmental attitudes are at all-time highs, these attitudes often fail to translate into environmentally sustainable behaviour. Much research has focused on exploring the predictors of, and barriers towards, sustainable behaviour. One such influence is social norms, though little research has focused specifically on the role of social learning. Our study focuses on three widely studied social learning strategies: payoff bias, prestige bias and conformity. Using a mixed methods approach, combining a questionnaire, vignette and behavioural measure with semi-structured interviews, we investigate the extent that Durham University students use social learning or are influenced by social information in relation to their sustainable behaviour. We also explore their general attitudes towards sustainability and perceived barriers of sustainable behaviour. We find some evidence that participants were influenced by their perceived norms regarding sustainability (indicating conformity) and that they factored the payoffs of sustainable behaviour into their decision making (indicating payoff bias) but there was little impact of high-status individuals (prestige bias). However, payoff, prestige and conformity framed vignettes all reduced the probability that participants selected the more sustainable voucher as a reward for participation and our interviews and questionnaire data provided greater indication for the use of other social learning strategies. More generally, the role of social influence was small compared to other factors such as personal attitude, cost, and systematic barriers.

6.2. Introduction

In the UK, since around 1972 and encouraged by EU membership, policy and dialogue surrounding sustainability has continued to develop (Christman, 2013). The importance of increasing sustainable behaviour cannot be understated. In 2019, 11,000 scientists supported claims of a "climate emergency" resulting largely from greenhouse gases due to human activities (Ripple et al., 2021). At global summits such as COP24, renewed interest from world leaders has encouraged more ambitious climate change targets to be introduced (COP24, 2018) although many argue that still not enough is being done (Ripple et al., 2021). More recently at COP26, further agreements have been made in attempts to keep global temperature increases below 1.5 Celsius, but many issues and barriers remain (Evans et al., 2021). Addressing climate issues at the global level is difficult: the costs associated with reducing emissions can be high and current day actions primarily benefit future generations (Carattini et al., 2019). It is also not possible to exclude those who do not make efforts to cut emissions from any global benefits that such actions produce (Carattini et al., 2019). This has led some to propose that the problem of sustainability can be understood as a social dilemma due to the incentive to free ride and the coordination it requires between individuals (Keohane & Victor, 2016). Sustainable behaviour can generally be defined as behaviour which does not compromise (or negatively impact) future generations on environmental, social or economic grounds (do Paço & Laurett, 2018).

Overwhelmingly, the primary source of sustained CO_2 emissions is the generation of energy which includes industry related emissions (such as manufacturing) and transportation (Ritchie & Roser, 2020). At the domestic level, households contribute by using energy (Stern, 2011), which accounts for around 10.9% of carbon emissions attributable to energy (Ritchie & Roser, 2020). An additional contribution is consumer demand for goods and services, both in terms of the ways and rate at which products are used within contemporary society (Tripathi & Singh, 2016). Despite this, the UK public's concern for the environment is at an all-time high (M. Smith, 2019), yet these positive attitudes often fail to translate into sustainable behaviours or purchases (Jaffe & Stavins, 1994; Tabi et al., 2014; White et al., 2019). This phenomenon, termed the “attitude behaviour gap”, has been widely acknowledged within sustainability research (Kollmuss & Agyeman, 2002; Nguyen et al., 2019).

6.2.1. Factors influencing sustainable behaviour

Many factors have been identified as important contributors to the adoption of sustainable practices across several populations and demographics (Faiers et al., 2007; Gifford & Nilsson, 2014; Tripathi & Singh, 2016; White et al., 2019) and individuals often show large variation in the extent to which environmental factors influence or motivate their purchase of sustainable products (Eberhart & Naderer, 2017). Personality traits such as agreeableness and openness (Hirsh, 2010; Hirsh & Dolderman, 2007) and holding biospheric (earth oriented) rather than egoistic (cost and benefit) values (de groot & Steg, 2008) have been shown to predict sustainable behaviour and environmental concern. Between countries, average agreeableness has been shown to correlate with the number of sustainable policies adopted by the country (Hirsh, 2014). Also, while it does not always predict sustainable behaviour (Mtutu & Thondhlana, 2016; Nguyen et al., 2019), holding a positive environmental attitude is often linked with sustainable behaviour (Best & Kneip, 2011; Gadenne et al., 2011) as is believing that one’s actions as a consumer will make a difference (Joshi & Rahman, 2019). Females on average report higher levels of sustainability than males but the link is inconsistent (Tripathi & Singh, 2016). Age is also inconsistent, where some studies find older people show greater environmental awareness and concern (D’Souza et al., 2007), others do not (N. Bhuian et al., 2014). A more consistent trend is associated with education, where university students reported greater environmental commitment, after controlling for other demographic variables, than the general population (Cotton & Alcock, 2013).

Further insights may be gleaned from direct interventions aiming to increase sustainable behaviour like recycling, purchasing sustainable products, or decreasing energy use. Providing environmental information can be effective at reducing household consumption (Delmas et al., 2013; Ek & Söderholm Patrik, 2010) but those which offer tailored information such as energy audits are more effective (Delmas et al., 2013). Emphasising further benefits associated with sustainability (for example, reduced pollution) can also be beneficial, particularly if the individual would especially benefit from it (Asensio & Delmas, 2015). Financial incentives are commonly used by many governments, for example to increase uptake of electric vehicles (Gov.uk, 2019) or encourage homeowners to undergo housing refits (Achtnicht, 2011). However, financial incentives can backfire if the offered rewards are too small (Handgraaf et al., 2013) or crowd out other potential motivators for sustainability (Bowles, 2008). Other strategies such as goal setting (Andor & Fels, 2018; Harding & Hsiaw, 2014) and providing detailed personal feedback on an individual's energy consumption (Vine et al., 2013) have also been shown to be effective in many cases.

Other research, often using interviews or questionnaires, instead focuses on the barriers individuals perceive to their sustainable behaviour. A commonly cited barrier is the price associated with sustainability (Acharya, 2019; Joshi & Rahman, 2015), as sustainable products tend to be more expensive than less sustainable products. Related to this, is the increased effort associated with sustainability, both in terms of the time investment required to act sustainably (Young et al., 2010) and the difficulties in acquiring products (Acharya, 2019). For example, a factor identified in the adoption of electric vehicles was having travel patterns that could accommodate them (Langbroek et al., 2016). Low levels of awareness or knowledge on how to be sustainable also inhibit individuals' sustainability (Axon, 2017; DEFRA, 2018; Young et al., 2010) as does being unmotivated to change one's behaviour (Thondhlana & Hlatshwayo, 2018). A study on consumers in India also found that individuals with high faith in technology also showed generally lower sustainable behaviour (Malodia & Bhatt, 2019).

6.2.2. Social norms

An area of research which has received specific attention is the role of social influence or peers on the adoption of sustainable behaviour (Andor & Fels, 2018). This may either be through direct copying of other individuals by social learning (Bandura, 1986) or by internalising social norms perceived in others (Cialdini et al., 1998) without the threat of punishment. Social norms are separated into descriptive (what do most people do) and injunctive (what do most people approve of) norms

(Cialdini et al., 1991). Research within sustainability has largely focused on the role of social norms (as opposed to social learning) and more specifically usually descriptive over injunctive norms. Some argue that social norms represent the most promising strategy available to reduce global greenhouse gas emissions (Carattini et al., 2019).

In general, social norms can have significant effects on sustainable behaviour (Yamin et al., 2019), despite many consumers believing the contrary (Nolan et al., 2008). In a systematic review, of the 24 studies considered, only 2 showed that normative information had no effect on sustainable behaviour (Andor & Fels, 2018). To date, the largest study of this kind was conducted in the US by OPOWER, an energy efficiency platform which provided energy usage reports to over 600,000 households (Allcott, 2011). Within the experiment, the reports compared their usage to similar households and suggested ways they could reduce their energy use. Overall, high usage households decreased their energy consumption after seeing descriptive norm comparisons with lower use households, while low use households showed smaller effects (Allcott, 2011). In contrast, a similar but smaller scale study found that descriptive norms produced a "boomerang effect" where both high and low use households would decrease and increase their energy use respectively to match the descriptive norm (Schultz et al., 2007). This was only avoided, by also providing households with injunctive norms, which ensured that low use households did not increase their usage (Schultz et al., 2007). This effect did not occur in the OPOWER study, which, although found no additional effect associated with injunctive norms, did accompany the descriptive norms with smiley faces which could be interpreted as an injunctive norm (Allcott, 2011). In other contexts, injunctive normative prompts were effective at encouraging drivers to switch off their car engines in queues at a busy level-crossing site (Mahmood et al., 2019), particularly if the prompt mentioned residents (an in group) versus visitors (Player et al., 2018). Hotel guests were more likely to reuse towels following a descriptive normative prompt than one which emphasised environmental benefits or the cost saving benefits (Goldstein et al., 2008).

Social norms are not effective in all cases. An intervention seeking to decrease energy use in a residential building found descriptive norms were less influential overall than other information such as personal feedback (Jain et al., 2012). A different study found that descriptive norms reduced households' energy use but not their gas consumption (Midden et al., 1983). Social norms may also be less effective when compared to interventions that make sustainability more convenient (Carlson, 2001) or when the expected normative behaviour is already very clear and widely accepted (Bergquist et al., 2020). Other demographic factors such as higher conservative political affiliation also reduced the effect that descriptive and injunctive norms had on reducing energy consumption (Costa & Kahn, 2013). The effectiveness of injunctive normative prompts may also depend on the environment, for example a normative request to switch off a light was less likely to be followed if the light was on when the individual entered the room (Bergquist & Nilsson, 2016). Likewise, littering was more likely in an already littered environment (Keuschnigg & Wolbring, 2015).

6.2.3. Social learning

In the context of sustainability, the direct copying of sustainable behaviour is considered much less frequently than more generalised normative influences. Though, there are examples where sustainable behaviour may have transmitted directly between individuals. Social network analysis showed peer feedback information was more effective at reducing energy use for individuals at the centre of the social network (with more social connections) who in turn were exposed to the information to a greater extent (Peschiera & Taylor, 2012). A choice experiment also found that participants were more likely to purchase sustainable chocolate with their earned points if they were informed that other participants had selected the sustainable option (Salazar et al., 2013). Within real life sustainability, there are several examples where individuals (or groups of individuals) facilitated the spread of beneficial information. In Switzerland, farmers produced a collaborative film to spread soil protection strategies to other farmers (Schneider et al., 2009) and participatory water management interventions that encourage participation and collaboration between stakeholders cause practices to spread through social learning (von Korff et al., 2012). A further example was seen in the transmission of lobstering practices in Maine, where there was some evidence that sustainable practices (such as escape vents for small lobsters) spread between fishers through social learning (Waring & Acheson, 2018).

Little research has considered the role of selective social learning strategies as considered within cultural evolution (R. Kendal et al., 2018; Laland, 2004). The consideration of social learning strategies arose from the assumption that unselective use of social learning is unlikely to be adaptive in most circumstances (Boyd & Richerson, 1995). Instead, individuals should use heuristics which influence when individuals use social learning, what kind of information they copy and from whom they acquire it (Laland, 2004; Morgan et al., 2012). Many different social learning strategies have been considered in previous work (R. Kendal et al., 2018). Of these, three have received the most widespread attention. Payoff bias, where individuals copy a trait or individual with a greater payoff than themselves (Schlag, 1998), prestige bias where individuals copy individuals who are generally successful and / or generally popular (Henrich & Gil-White, 2001) and conformity where individuals adopt the trait displayed by the majority (Henrich & Boyd, 1998). To date, little to no research has specifically investigated payoff bias or prestige bias with respect to sustainability. Conformity is functionally similar to descriptive social norms and so will not be considered further in this section.

A payoff bias could motivate sustainable behaviour through the potential monetary savings offered by sustainable behaviour (as is the case with reducing energy use). Direct financial rewards for reducing the use of energy are effective in some cases (Jain et al., 2012). Undertaking housing refits are also predictable based on expected monetary savings (Frondel & Vance, 2013) and if the government offers generous tax rebates (Durham et al., 1988). Alternatively, sustainable behaviour may offer intrinsic benefits, promoting a “do good, feel good” mentality, similar to the warm glow effect (Andreoni, 1990). Two such examples are interventions encouraging the recycling of bottles, where emotional messages highlighting expected positive feelings towards recycling outperformed injunctive normative messages (Bergquist et al., 2020; Neumann, 2019). Conversely, a payoff bias could easily erode sustainable behaviour if participants focused on the greater cost or inconvenience of sustainable behaviours, rather than the potential benefits. If sustainability is considered as a prisoner’s dilemma (Keohane & Victor, 2016; Rapoport, 1974) payoff learning is predicted to result in the extinction of cooperative behaviour in most circumstances (Doebeli & Hauert, 2005a).

Under a prestige bias, models are copied in a non-specific domain general way based on their status or general skill (Henrich & Gil-White, 2001; Jiménez & Mesoudi, 2019). As such, it is plausible that high-status individuals (such as celebrities) may influence the sustainable behaviour of their followers. One example is Leonardo DiCaprio who promotes pro-environmental attitudes but is famous for acting (WWF, 2021). In addition to cases like Leonardo DiCaprio, other individuals specifically famous for their pro-environmental attitudes such as David Attenborough or Greta Thunberg may also be influential for the spread of pro-environmental behaviour. There has been only very limited attention granted to the influence of high-status individuals on sustainable behaviour. A survey and resulting model found that the intention to purchase luxury sustainable products was increased with celebrity endorsement (Cuomo et al., 2019) and the uptake of websites which offer rental clothes was also influenced by Instagram "micro celebrities" with large followings that promote it (Shrivastava et al., 2020). A recent study comprised of a representative US sample found familiarity with Greta Thunberg predicted intention to take actions to reduce global warming (Sabherwal et al., 2021) which provides a more direct indication of a prestige bias. Outside of sustainability, the link between prestige and real-life cooperative behaviour is inconsistent. For example, where prestige conferred upon team leaders was correlated with the team's cooperation in the royal navy (Offord et al., 2016, 2019), celebrity endorsement did not increase volunteering in students (John et al., 2019) or the likelihood that young voters voted in general elections (N. T. Wood & Herbst, 2007) .

In a more general sense, there is evidence that sustainable behaviour may spread more efficiently via a single individual than through multiple. Greater uptake of recycling in a block of flats was found when information was spread through an elected leader than through information packages left at residents' doors (Burn, 1991). In a workplace setting, a survey found a strong correlation between the sustainable behaviour of team leaders and other employees in the company (Robertson & Barling, 2013). A series of interviews also identified that a wetlands conservation project in Sweden had largely begun, been managed by, and spread through a single individual (Olsson et al., 2004).

6.2.4. The present research

Our study has two aims. Primarily, we seek to expand and contribute to the growing literature seeking to apply cultural evolution to sustainability (Brooks et al., 2018). Specifically, we will assess the extent to which individuals use, or are influenced by, a payoff bias (Schlag, 1998), prestige bias (Henrich & Gil-White, 2001) or conformity (Henrich & Boyd, 1998) within their sustainable decision making. We also consider the ways in which the behaviour or attitudes of others more generally affects their behaviour. Secondarily, we will also explore participants perceptions of other factors that are relevant for their sustainable behaviour. Finally, we assess participant's understanding of and normative beliefs towards sustainability and any barriers they perceive to their sustainable behaviour.

The population we focus on for this study is Durham University students. While the overuse of student samples continues to plague many studies within psychology, particularly those from WEIRD backgrounds (Henrich, Heine, et al., 2010), as achieving a more conventionally representative sample was not possible for this study, we chose to specifically recruit from a single population. Focusing on a restricted population does offer some benefits. There are many factors which influence sustainable behaviour, several of which are demographic variables (Tripathi & Singh, 2016). Focusing on participants from a restricted population whom share some of their demographic backgrounds can help control for some of this variability. In general, students are usually well informed on environmental issues and hold positive attitudes towards sustainable behaviour (Emanuel & Adams, 2011; Levine & Strube, 2012) and other research has identified that those who have completed higher education levels typically engage in more sustainable behaviour (Cotton & Alcock, 2013). Nevertheless, we do not intend to generalise our results to other populations. Instead, we seek to evaluate our research questions within this population (which we assume our sample is representative of) where our findings may then be compared with other studies comprised of a different sample investigating similar questions (Smith, 2018).

In this study, a mixed methods approach is employed, which is especially appropriate for a complex and multidimensional issue like sustainability. This allows specific variables of interest to be investigated and participants to express their understanding of the issue in their own words (Eagly & Riger, 2014). As such, our study is split into two parts. Study one involves an online questionnaire including vignettes and a behavioural measure similar to Salazar et al., (2013). The questionnaire was designed to assess participants perceptions of sustainable behaviour, the extent to which they believe others influence their behaviour and the barriers they perceive to their sustainable behaviour. The vignettes and behavioural measure offer an experimental test of the effect of information framed as payoff, prestige, and conformity information on participants' sustainable behaviour. To measure sustainable behaviour, participants were offered the choice to enter a raffle for a voucher redeemable at a sustainable or less sustainable online clothing retailer. In study two, 7 individuals from the questionnaire were recruited to a semi-structured interview where the general themes addressed in the questionnaire could be explored in greater detail. Data were analysed using thematic analysis (Braun & Clarke, 2006, 2019), where recurring ideas expressed by participants were grouped into overarching themes.

6.2.5. Research questions

As this study is primarily exploratory in nature and we are not necessarily seeking to test any specific hypotheses, we instead summarise the goals of the study through two overarching research questions (RQ), which we evaluate in both parts of our study.

RQ1) To what extent are Durham University students influenced by others within their sustainable decision making? – We focus particularly on exploring the role of payoff based, prestige based and conformity-based copying (which we also assess using the experimental manipulation). However, the goal of the study is not restricted to these possibilities and will assess the influence of others on sustainable behaviour in the broad sense.

RQ2) What are the attitudes and perceptions towards sustainability of Durham University students? –Using the questionnaire and the interviews, we seek to explore participants attitudes towards sustainability, particularly in terms of their social norms and the barriers they perceive within their sustainable behaviour.

6.3. Study 1 Methods

The questionnaire addresses our research questions in two ways. We collect self-report data, where we ask participants to rate the influence of select individuals on their sustainable behaviour (RQ1) and their personal norms regarding sustainable behaviour (RQ2). We also investigate the barriers participants perceive to their sustainable behaviour (RQ2). We additionally use Principal components analysis (PCA) to evaluate the extent to which participant's responses are explainable by underlying factors. We also include an experimental design and a behavioural outcome measure to assess the influence of payoff, conformity, or prestige framed vignettes on participant's sustainable decision (RQ1).

6.3.1. Participants

Durham University students completed a 15-minute online questionnaire. Participants were recruited from a variety of sources including student bulletins or pages, word of mouth and a Durham University participant recruitment system. Participants were not paid for their participation but had the option to enter a raffle for a clothing voucher (all but 26 participants opted to do so). 231 participants began the questionnaire, but after excluding those who did not complete the questions used in PCA, 182 participants were retained in the data and comprised of 30 Males (16%), 147 (81%) Females, 3 Non-binary (2%), 2 (1%) Prefer not to say with a median age of 20 years (IQR = 3). Students were from a range of departments and years of study. The four most common departments were: Psychology (N=51), Anthropology (N=34), Biological Sciences (N=14) and Geography (N=13).

6.3.2. Materials

We developed questionnaire materials by adapting established scales, drawing on findings from previous studies and theoretical constructs (questionnaire is provided in SI 1). We used PCA to evaluate the extent to which the questionnaire responses loaded onto underlying factors, which then produced single summary values for use in a GLM. Below we describe each questionnaire block. Unless otherwise stated, participants provided their response on Likert scales from 1 (*strongly disagree*) to 4 (*strongly agree*). The scale was fully continuous, such that participants could give a value up to 2 decimal places.

6.3.2.1. Level of sustainable behaviour

Participants rated how often they engaged with nine different sustainable behaviours using a scale from 1 (never) to 4 (always). Examples included: *"Only boil the amount of water I need in the kettle"* and *"Use a reusable coffee cup".* These questions were adapted from relevant items identified in previous research on students (Armel et al., 2011; Casey & Scott, 2006; F. G. Kaiser & Biel, 2000; Walton & Austin, 2011).

6.3.2.2. Personal attitudes towards sustainable behaviour

To measure a selection of general environmental attitudes held by participants we developed six items motivated by a review of the literature and loosely inspired by an established scale (Minton & Rose, 1997). For example: *"I have a good awareness of the environmental and societal benefits of sustainable behaviour"* was based on studies finding knowledge as a barrier to sustainability(Axon, 2017; Murphy, 2014; Young et al., 2010) and *"I am likely to join in and behave sustainably if I see others around me doing the same"* to evaluate participants beliefs that they are influenced by others, which is often underreported (Nolan et al., 2008).

6.3.2.3. Barriers to sustainable behaviour

Participants rated the extent that ten barriers prevented them from behaving sustainably in their day-to-day life. These included: *"I am forgetful"*, *"Others around me are not behaving sustainably"* and *"I feel like my behaviour will not make any difference".* These items were based on barriers identified by previous research on students (Thondhlana & Hlatshwayo, 2018), but also influenced broadly by other studies focused on barriers to sustainability (Acharya, 2019; Young et al., 2010). These questions also provided some indication of the extent to which payoffs associated with sustainability were relevant for participants' decisions (Schlag, 1999). Participants were also offered a free text box, where they could state any further barrier which was not listed, though few participants offered a response (appendix 6.2).

6.3.2.4. External environmental norms

To measure injunctive norms regarding sustainability (Cialdini et al., 1998), participants rated the extent that their friends, family, Durham University students, and the general public felt a personal moral obligation to behave sustainably. This also provides insight into the role of conformity (Boyd & Richerson, 1985) once this is correlated with participant's sustainable behaviour.

6.3.2.5. Environmental concern

To measure environmental concern, we used six items taken from the established environmental concern scale (Minton & Rose, 1997). This scale evaluates an individual's personally held attitude towards the environment. Questions relevant to students were included while those relating to pollution were dropped as it was beyond the scope of this study. Examples of questions were: *"We are not doing enough to save scarce materials from being used up"* and *"non-recyclable containers should be taxed to reduce waste"*.

6.3.2.6. Sources of social influence

To investigate sources of social influence, participants were asked to rate how likely different people (or groups of people) would be to influence their sustainable behaviour from 1 (*very unlikely*) to 4 (*very likely*). Friends and family were hypothesised as sources of influence for conformity, but we note that we cannot assess conformity in the strictest sense (disproportionate copying of common traits) as originally defined by (Boyd & Richerson, 1985). Prestige is defined as someone of high status or who is culturally popular (Henrich & Gil-White, 2001), but very little work has investigated prestige in relation to sustainability, particularly outside of celebrities (Cuomo et al., 2019; Sabherwal et al., 2021; Shrivastava et al., 2020). We offered several alternatives including *climate activists* and *local individuals with large social media followings* as well as *celebrities*. A payoff bias would rarely predict sustainable behaviour as it is usually more costly than less sustainable behaviour, though it is possible participants confer benefits on those who behave sustainably. To assess this possibility, participants were asked to rate the extent that they respected those who behaved sustainably.

6.3.3. Procedure

Participants completed the questionnaire on their own device and accessed it using an anonymous link or a QR code. Participants first read and accepted an information sheet and consent form. They were then presented the question blocks. Participants completed all parts of the questionnaire but could abstain from answering any question. Participants were then presented with one of three randomly selected vignettes of roughly equal lengths or no vignette in the control condition. The vignettes described a fictional individual Alex and emphasised information pertaining to conformity, payoff, or prestige biases (full vignettes are given in appendix 6.1). To avoid any influence of gender, gender neutral pronouns were used throughout (they/them). For Prestige, Alex was described as an award-winning actor who supported efforts to reduce climate change. Conformity quoted a survey suggesting 70% of the UK population believe greenhouse gases are to blame for temperature increases and described Alex’s neighbourhood as responding by becoming greener. Payoff described

Alex noticing the various personal benefits associated with sustainable behaviour including cheaper bills, personal satisfaction, and respect from co-workers.

After completing a short series of demographic questions (including age, gender, whether they consume animal products, and their political affiliation) participants were offered the chance to enter into a prize draw for a clothing voucher. They were offered a choice between vouchers for online retailers ASOS (less sustainable option) or Rapanui (more sustainable option). While both retailers promote sustainable practices on their website, Rapanui produces their clothes from exclusively organic or recycled material and uses only renewable energy in their production lines but are more expensive on average than ASOS. This choice was offered to act as an outcome variable alongside the vignettes to test the effect of conformity, prestige, and payoff biases on real-world sustainable behaviour.

6.3.4. Data analysis

All analyses were conducted in RStudio version 4.1.2 (R Core Team, 2021). Data and analysis scripts are available in SI 2 (https://osf.io/frpnu/?view_only=4bfbaccdf7d142fa87093f5aa5bb4ad8). Besides reporting participants' responses descriptively, the data were explored using PCA and GLMs. PCA was used to identify underlying factors that may explain correlations between questions. The number of factors was determined using parallel analysis (Horn, 1965) as opposed to the Kaiser criterion (H. F. Kaiser, 1960) as parallel analysis generally outperforms the Kaiser criterion in simulation studies (Braeken & van Assen, 2017) and is more conservative. Factor scores were generated for each participant based on their responses. The factor scores where then used alongside the experimental manipulation and several other variables in Bayesian logistic regression models to predict choosing the sustainable voucher. We use descriptive statistics and the Bayesian models to assess the influence of others on their sustainable behaviour (RQ1) and use PCA and descriptive statistics to assess participants' attitudes towards sustainable behaviour (RQ2).

Bayesian models were run with 4 chains of 500 warmup samples and 1500 samples for inference using the *brms* package (Bürkner, 2017) and Markov Chain Monte Carlo to construct the posterior distribution. Four models were fit to the data to evaluate the impact of the experimental manipulation and the other independent variables on participants choices. “Control” contained a single intercept parameter and served as a baseline model for comparison. “Condition” included a varying intercept for the experimental condition and “Responses” contained only the parameters associated with the items and factor scores from the questionnaire (henceforth, response variables). “Condition+Responses*”* contained both the condition and response parameters. The mathematical definition of the “Condition+Responses” model is shown below.

$$Choosing\ sustainable\ option \sim Binomial(1, p)$$

$$logit(p) = \alpha_{condition} + \beta_1 * Vegan\ or\ Vegeterian + \beta_2 * Conservatism + \beta_3 * Female + \beta_4 * Perceived\ Barrier + \beta_5 * Acting\ sustainably + \beta_6 * Perceived\ Duty + \beta_7 * Student\ norms + \beta_8 * Family\ norms + \beta_9 * \text{Friends norms}$$

$$a_{condition} \sim Normal\ (0, 1.5)$$

$$\beta_{1:9} \sim Normal\ (0,2)$$

The predictor variables are constructed as follows. Vegan or vegetarian is a binary variable, where a value of 1 indicates they are vegetarian or vegan (these groups were combined due to the small number of vegans in the sample). Conservatism is a score from 1-4 indicating how politically conservative the participant is. Female is a binary variable where a value of 1 indicates they self-identified as Female. Because of a substantial gender bias in the sample, self-identified males were pooled with the participants who identified as non-binary or those who withheld the information. Perceived Barrier, Acting Sustainably and Perceived duty are the factor scores identified from PCA (section 6.4.2.). Student norms, Family norms and Friends norms are the scores (from 1 – 4) provided by participants regarding the strength of the environmental injunctive norms they perceived in these individuals. The response variable, choosing sustainable option, was coded as a 1 if they selected the Rapanui voucher and 0 if they selected the ASOS voucher.

Data were excluded if the data were incomplete for the questions used in the PCA or if the participant had identified as a non-student. Cases where participants had missing data for any of the variables used in the Bayesian model but otherwise complete data were retained for analysis. In this case, imputation was used to generate 20 datasets which simulated this missing data. Then, equivalent Bayesian models were fit to these datasets and the posterior distribution was pooled across them (Bürkner, 2021).

The effect of the response variables and experimental condition were evaluated using model comparison, inspecting the posterior distribution and plotting model predictions. Using model comparison, the relative fit of competing models which include different combination of predictor variables may be assessed on their out of sample deviance using WAIC (widely applicable information criteria) and applying a penalty based on the number of parameters (Kruschke, 2015; McElreath, 2020). However, model comparison is inappropriate for pooled model results as model fitting occurs across different datasets. As such, for the purposes of model comparison, models used a single imputed dataset. The best fitting model was then refit using the imputed datasets and the posterior distribution was pooled across these models. The impact of the response predictors was evaluated from this pooled model by plotting the posterior distributions of the β values and considering the proportion of the posterior that is in the same direction as the mean (as in Martin et al., 2021) and producing counterfactual predictive plots. This provides an estimate of the plausibility of a positive or negative effect.

6.4. Study 1 Results

6.4.1. Descriptive analysis

Here we provide descriptive statistics for some of the items rated by participants. Included items are those not used in PCA or where a descriptive breakdown beyond overall factor scores helps answer our research questions. Note that in every case, a high score indicates high agreement with the question item.

6.4.1.1. Participant's perceptions of barriers

Table 6.1 shows the descriptive statistics for each of the perceived barrier items. Overall, the barriers identified as the most common were the costs of being sustainable (*M* = 2.55, *SD* = 0.80), being forgetful (*M* = 2.53, *SD* = 0.80) and the lack of others around them being sustainable (*M* = 2.53, *SD* = 0.75). Conversely, feeling someone else will take care of it (*M* = 1.74, *SD* = 0.64), preferring less sustainable alternatives (*M* = 1.80, *SD* = 0.72) and lack of knowledge (*M* = 1.88, *SD* = 0.66), were rated as the least common barriers.

Table 6.1. *Means and standard deviations of participant's responses to the question "How much do you agree / disagree that the followings aspects prevent you from behaving sustainably in your day-to-day life?".*

Item	*Mean (SD) (1 – 5)*
I cannot afford the additional costs of being sustainable	2.55 (0.80)
Others around me are not behaving sustainably	2.53 (0.80)
I am forgetful	2.53 (0.75)
It is inconvenient	2.39 (0.73)
I feel like my behaviour will not make any difference	2.23 (0.78)
It's too much effort	2.18 (0.67)
I do not have control over my sustainability	2.06 (0.70)
I do not know how to behave more sustainably	1.88 (0.66)
I prefer less sustainable alternatives	1.80 (0.72)
I feel that someone else will take care of it	1.74 (0.64)

6.4.1.2. Evaluating the sources of social influence

For the sources of social influence measure (Table 6.2), the highest rated option was scientists (*M* = 3.53, *SD* = 0.46), followed by friends (*M* = 3.28, *SD* = 0.56), and family (*M* = 3.23, *SD* = 0.65). The lowest rated were celebrities (*M* = 2.11, *SD* = 0.80) and individuals with large social media followings (*M* = 2.25, *SD* = 0.82). Standard deviations were lower for scientists and friends than the other items, indicating more consistency in participant's answers. Participants offered generally high agreement with the statement "*I am likely to join in and behave sustainably if I see others around me doing the same*" (*M* = 3.36, *SD* = 0.60), suggesting they perceived some effect of others on their behaviour. They also provided strong agreement with the statement "*I respect those who behave sustainably*" (*M* = 3.81, *SD* = 0.34), which suggests that those who are sustainable may gain other benefits in the form of status.

Table 6.4. *Means and standard deviations of participant's responses to the question "How likely/unlikely would the following people be to influence your decision to behave sustainably?".*

Response	*Mean (SD) (1 – 5)*
Scientists	3.53 (0.46)
My friends	3.28 (0.56)
My family	3.23 (0.65)
Climate activists	2.93 (0.80)
Local individuals with a large social media following	2.25 (0.82)
Celebrities	2.11 (0.80)

Descriptive analyses of the external environmental norm items (Table 6.3) demonstrated that participants believed that their friends (*M* = 2.93, *SD* = 0.58), family (*M* = 2.82, *SD* = 0.74) and fellow students (*M* = 2.74, *SD* = 0.59) all felt a stronger environmental norm than the general public (*M* = 2.36, *SD* = 0.70).

Table 6.5. *Means and standard deviations of participant's responses to the question "How much do you agree / disagree that the following groups of people feel a personal moral obligation to do everything they can to protect the environment?".*

Question	*Mean (SD) (1 – 5)*
My friends	2.93 (0.58)
My family	2.82 (0.74)
Durham university students	2.74 (0.59)
The general public	2.36 (0.70)

6.4.2. Principal components analysis

PCA was conducted on 30 items, which includes the questions regarding participants' level of sustainable behaviour, personal attitudes towards sustainable behaviour, barriers to sustainable behaviour, and environmental concern. The items related to external social norms and sources of social influence were not included because these questions were designed to assess the strength of influence for a given item independently of one another. There was also no reason to assume a-priori that if individuals were influenced by their friends, they would also be influenced by other groups of individuals.

The Kaiser-Meyer-Olkin measure was 0.91 overall and all KMO values for items were above the recommended .5 cut-off (Kaiser, 1974). Additionally, Bartletts test of sphericity (χ^2 (465) = 1795.835, $p < 0.001$) indicated sufficient correlation between items for PCA and the determinant was above 0.00001 which suggested no multicollinearity. As such, the data were suitable for PCA. Oblique factor rotation was used (oblimin) to allow the components to be correlated.

Initial analysis obtained eigenvalues for each individual component. Then, to determine the number of components to retain, parallel analysis was used to generate eigenvalues from simulated data to compare to the eigenvalues from the data (Horn, 1965). Any components with an eigenvalue of less than the simulated value was dropped. This supported retaining three components in the final analysis. At this stage, the question item *"Reuse bags"* from the level of sustainable behaviour question block was identified to load poorly onto the three components, so the analysis was repeated after dropping this item.

Although some items load onto multiple components, the patterns of item loadings suggest that component 1 represented a **perceived concern and duty** towards the environment, component 2 represented a **perception of barriers to sustainability**, and component 3 represented a **willingness to behave sustainably**. Because oblique factor rotation was used, these components were weakly correlated. Perceived concern was negatively correlated with perceptions of barriers (r = -0.22) but positively correlated with a willingness to behave sustainably (r = 0.22) and perceptions of barriers was negatively correlated with a willingness to behave sustainably (r = -0.28). Factor loadings between components are shown in Table 6.4. Component scores were generated for each participant to be used in the Bayesian models. Correlations between the factor scores and predictors included in the Bayesian models are given in appendix 6.3.

Table 6.6. *Item loadings on the three components after oblique rotation (N = 182). Items in bold are those that load above 0.3.*

Item	*Perceived concern and duty towards the environment*	*Perceived barriers to sustainability*	*Willingness to behave sustainably*
We are not doing enough to save scarce materials from being used up	**0.83**	0.03	-0.11
The government should devote more money toward supporting conservation and environmental programs	**0.81**	0.01	0.03
Manufacturers should be required to use recycled materials in their operations whenever possible	**0.74**	-0.02	-0.01
Natural resources must be preserved	**0.70**	-0.05	-0.06
Non-recyclable containers should be taxed to reduce waste	**0.63**	-0.13	-0.04
I want to behave more sustainably than I currently am	**0.53**	0.06	0.28
Consumers should be interested in the environmental consequences of the products they purchase	**0.51**	0.01	**0.35**
We are not doing enough to save scarce materials from being used up	**0.83**	0.03	-0.11
I feel that I have a moral obligation to do everything I can to protect the environment	**0.42**	-0.23	**0.37**
I am likely to join in and behave sustainably if I see others around me are doing the same	**0.35**	0.25	0.29
It's too much effort	0.18	**0.72**	-0.13
It is inconvenient	-0.05	**0.66**	0.00
I do not know how to behave more sustainably	-0.07	**0.63**	0.07
I cannot afford the additional costs of being sustainable	0.09	**0.55**	0.13
I prefer less sustainable alternatives (for example, plastic wrapped fruit / disposable coffee cups or using artificial lights)	**-0.36**	**0.48**	0.02
Others around me are not behaving sustainably	0.02	**0.47**	**0.45**
I feel like my behaviour will not make any difference	-0.07	**0.47**	-0.12
I have a good understanding of how I can behave more sustainably	0.20	**-0.46**	0.26
Recycle items that can be recycled (e.g. paper, drink bottles and cardboard)	0.21	**-0.45**	0.04
I do not have control over my sustainability	0.09	**0.45**	0.10

I am forgetful	0.05	**0.43**	0.01
I feel that someone else will take care of it	-0.24	**0.42**	-0.14
I have a good awareness of the environmental and societal benefits of sustainable behaviour	0.21	**-0.40**	**0.30**
I feel that behaving more sustainably is pointless	-0.29	**0.31**	-0.15
Turn off the lights when I leave the room	-0.04	0.04	**0.58**
Use a reusable coffee cup	-0.04	-0.03	**0.50**
Switch off or put my computer to sleep when not in use	-0.18	-0.09	**0.46**
Wash clothes using the coldest setting	-0.08	-0.15	**0.46**
Only boil the amount of water I need in the kettle	**-0.37**	-0.22	**0.45**
Avoid products with excessive packaging	0.22	-0.19	**0.44**
Rely on sunlight rather than having the light on	0.17	-0.12	**0.39**
Eigenvalues	4.56	4.07	2.69
% of variance	15	14	9

6.4.3. Bayesian model comparison

To evaluate the influence of experimental condition on choosing the sustainable voucher, the three factor scores (duty, barriers and willingness), the other demographic predictors (female, conservatism, vegan / vegetarian) and participants perceived norms (Durham students, friends and family) were included in Bayesian GLMs. Model comparison using WAIC was conducted between models which contained different combinations of these variables. A single dataset was imputed which simulated the missing values in choosing sustainable option (26 missing), conservatism (4 missing), vegan/vegetarian (1 missing), Durham student norms (3 missing) and Friends norms (1 missing), and models were then fit to this dataset. The WAIC values and associated model weights are shown in Table 6.5. All continuous variables were standardised.

Table 6.5. *WAIC values and model weights for models evaluating the impact of including different predictor variables. SE shows the standard error in the WAIC estimate while standard error difference shows the standard error in the difference between each model and the model with the lowest WAIC value. Responses refers to the model including the demographic and factor score predictors.*

Model	*WAIC*	*SE*	*Standard error difference*	*Weight*
Condition+Responses	202.3	16.8	0	1
Responses	218.2	14.7	5.2	0
Condition	233.6	9.5	7.3	0
Control	243.6	6.5	8.1	0

Model comparison offered unambiguous support for the Condition+Responses model, which had the lowest WAIC score (202.3) and occupied all of the model weight. The inclusion of both the experimental condition and the other predictor variables improved out of sample predictions in relation to the intercept only model (Control: WAIC = 243.6) and models which included only experimental condition (Condition: WAIC = 233.6) or the demographic and factor score predictors (Responses: WAIC = 218.2). Therefore, the Condition+Responses model was refit to all 20 imputed datasets and the posterior distribution was averaged between them. We note that the Responses model is also favoured over the Condition model by WAIC, which suggests there is a larger effect associated with the questionnaire responses than the experimental condition.

6.4.3.1 Predicted probabilities across experimental condition

Predictions generated from the averaged Condition+Responses model are shown in Figure 6.1. Overall, the probability that participants selected the sustainable voucher was highest in the control condition (M = 0.88; 95% PI = 0.73, 0.97) where they did not read a vignette. Predicted probabilities of selecting the sustainable voucher were similar between the prestige (M = 0.70; 95% PI = 0.46, 0.88) and conformity (M = 0.62; 95% PI = 0.38, 0.83) conditions but lowest in the payoff condition (M = 0.45; 95% PI = 0.20, 0.71).

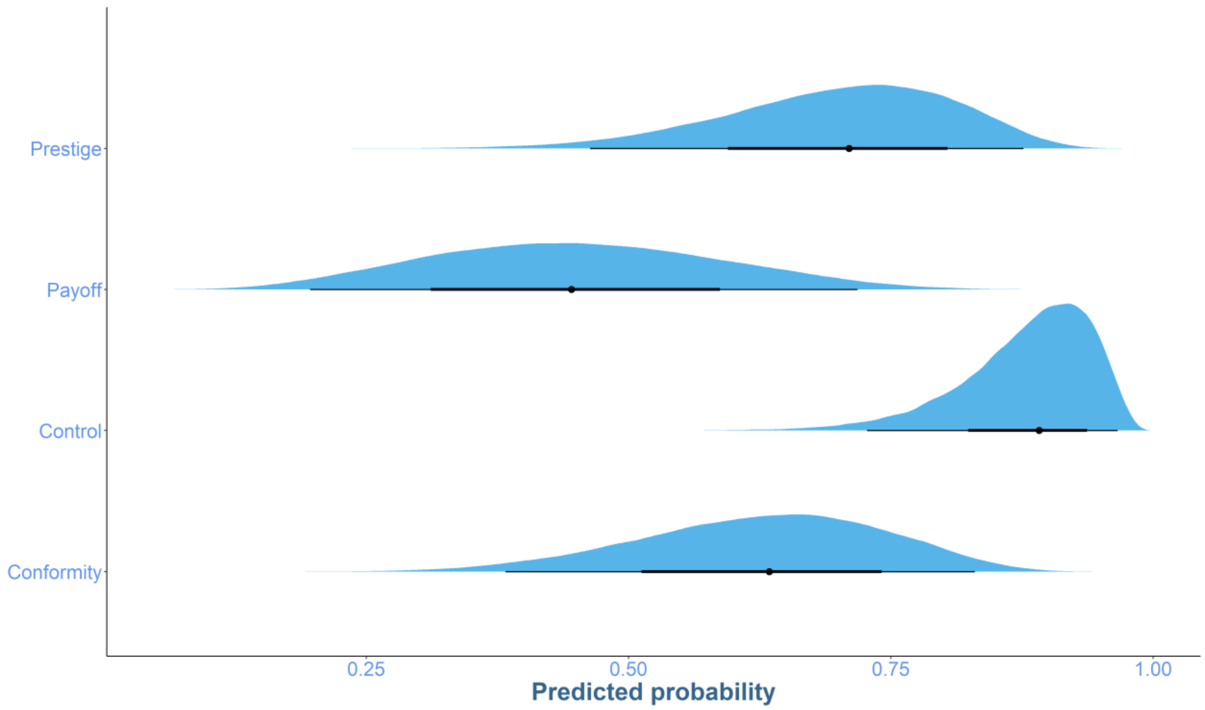


Figure 6.1. *10000 predicted probabilities of selecting the sustainable voucher drawn from the pooled Condition+Responses model. Points show the mean probability per condition and lines show the 68% and 95% PI across the posterior distribution. All other predictors are set at 0 (their average).*

6.4.3.2. Parameter estimates for the demographic variables and the factor scores

To assess the impact of the factor scores and demographic variables on the probability of selecting the sustainable voucher, parameter estimates were extracted from the Condition+Responses model and shown in Figure 6.2. Being vegan or vegetarian and high scores on the perceived concern and acting sustainably components most strongly positively predicted selecting the sustainable voucher. Being female and political conservatism most strongly negatively predicted selecting the sustainable voucher. The posterior distributions of the remaining predictors were not reliably above or below 0.

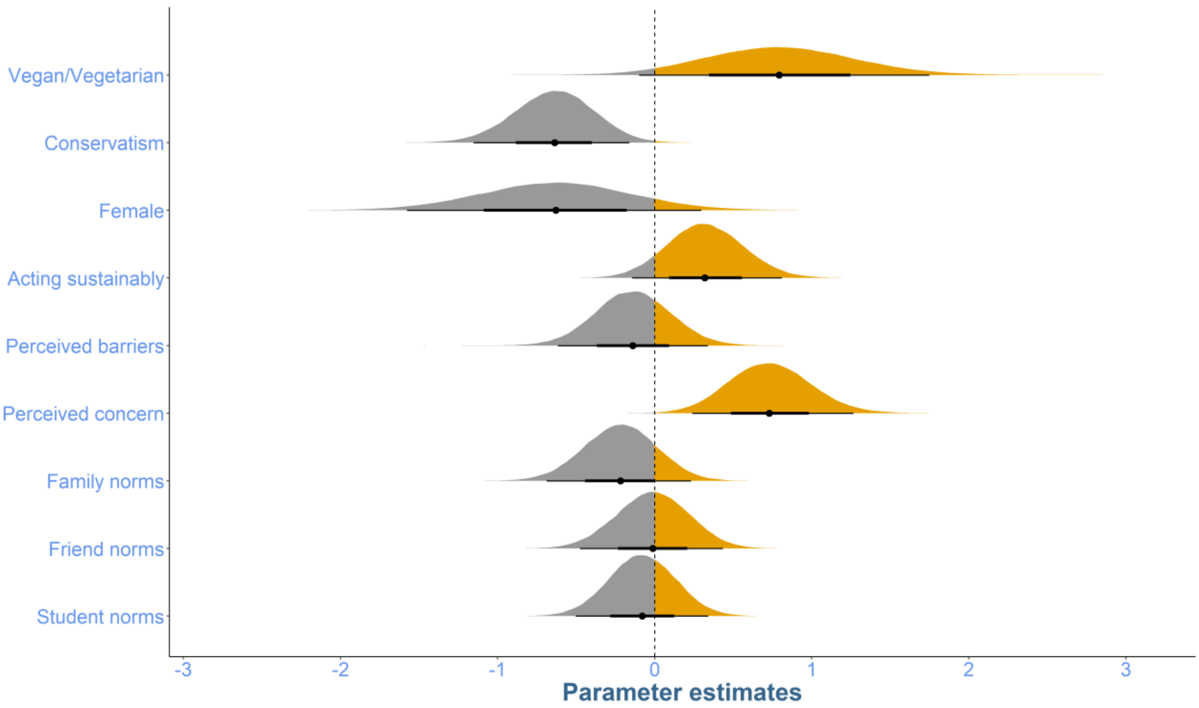


Figure 6.2. *10000 samples of the posterior distribution of condition and response variable parameters from the Condition+Responses model. Yellow shading indicates regions of the distribution above 0 and grey shading indicates regions of the distribution below 0. Points and solid lines show the mean and 68% and 95% PI for each parameter.*

Table 6.6 shows the mean parameter estimates, 95% prediction intervals (PI) and the proportion of the posterior that is on the same side of 0 as the mean of the response predictors included in the "Condition+Response" model. This provides strong support for a positive effect of the perceived concern component and being vegan or vegetarian and a negative effect of being highly politically conservative. The evidence for a negative effect of being female on acting sustainably is also strong, but less certain than the other variables. The evidence is weakest for perceived barriers, where 63.81% of the posterior is negative. Of the three social norm measures included (friends, family and students), all were weakly negative, but family showed the largest effect, where 82.75% of the posterior was negative. Both friends (51.96%) and student (64.44%) overlapped 0 considerably.

Table 6.7. Mean estimates and 95% PI of the posterior distribution generated from the Condition+Responses model. Also shown is the percentage of the posterior that lies on the same side of 0 as the mean (given to 2 decimal places).

Parameter	*Mean estimate (95% PI)*	*% in direction of mean*
Vegan/Vegetarian	0.80 (-0.10;1.7)	95.86%
Conservatism	-0.64 (-1.15; -0.2)	99.60%
Female	-0.63 (-1.58; 0.3)	90.77%
Acting sustainably	0.32 (-0.14; 0.8)	91.05%
Perceived barriers	-0.14 (-0.62;0.3)	71.81%
Perceived concern	0.74 (0.24; 1.3)	99.85%
Family norms	-0.22 (-0.69; 0.2)	82.75%
Friend norms	-0.01 (-0.47;0.4)	51.96%
Student norms	-0.08 (-0.50; 0.3)	64.44%

6.5. Study 1 Discussion

To address our first research question (RQ1: to what extent are Durham University students influenced by others within their sustainable decision making), we asked participants to rate the relative influence of different individuals on their sustainable behaviour and their perceptions of other individuals' injunctive norms around sustainability. We also experimentally manipulated the salience of three social learning strategies (payoff, prestige, conformity) using vignettes and assessed their effects using a behavioural measure. We found some evidence of conformity, indirect evidence of payoff bias but little evidence of prestige bias. Participants were influenced to some degree by the behaviour of others around them (conformity) and perceived cost and inconvenience as barriers to their sustainable behaviour (payoff). However, the probability of selecting the sustainable voucher was reduced by all three vignette conditions relative to the control condition and was not strongly related to the injunctive norms of participants' family, friends, or fellow students.

To address our second research question (RQ2: What are the attitudes and perceptions towards sustainability of Durham University students?) we asked participants to rate the primary barriers they perceived to their sustainable behaviour and complete a battery of questions assessing their environmental attitudes. We also employed PCA to identify underlying components. PCA revealed three key components: (1) perceived environmental concern, (2) perceived barriers to sustainability and (3) willingness to behave sustainably. Participants generally held strong injunctive environmental norms and felt their primary barriers were the costs and extra effort associated with sustainability and being discouraged by the lack of sustainability in others.

Firstly, considering the influence of others, and particularly the role of social learning biases (R. Kendal et al., 2018), participants reported scientists as the most likely group of individuals to influence their sustainable decision making. Scientists may have been influential due to their status, implying a prestige bias (Henrich & Gil-White, 2001; Jiménez & Mesoudi, 2019) or their expertise, implying a knowledge-based bias (R. Kendal et al., 2015; L. A. Wood et al., 2015). We argue a knowledge bias is more likely than a prestige bias where individuals are influential beyond their domain of expertise (Henrich & Gil-White, 2001). Though the questionnaire did not specify the kind of scientist, a climate scientist's expertise would be within the domain of sustainability, so a heuristic strategy like prestige would not be necessary. Furthermore, the other prestigious individuals that participants were asked about (celebrities and those with a large social media following) were rated as the least likely source of influence, suggesting participants were not using a generalised prestige bias to select their information source. This is somewhat consistent with previous findings regarding prestige that are restricted to investigating behavioural intent. Celebrity endorsement was shown to increase the purchase intention of luxury sustainable products (Cuomo et al., 2019) but the intention to go vegan only influenced those who already did not eat meat (Phua et al., 2020). Outside of sustainability, though theoretical models predict prestige can sustain cooperative behaviour (Henrich et al., 2015), many real-life cooperative behaviours like voting (N. T. Wood & Herbst, 2007) and volunteering (John et al., 2019) are not influenced by celebrity endorsement.

The role of a payoff bias may be indirectly inferred from the responses participants provided to the barrier items. Participants rated the cost of being sustainable as their top barrier, which has been shown in previous studies (Horhota et al., 2014; Joshi & Rahman, 2015). This is unsurprising, given sustainable products (like renewable energy (TheGreenAge, 2019)) are often more expensive than less sustainable products and students may lack the resources to afford more sustainable products. Other highly rated barriers were being forgetful and inconvenience which implies that sustainability requires additional effort and systematic barriers may be influencing participant's behaviour. This too is supported by previous research, showing people feel sustainability is difficult to integrate into busy schedules (Axon, 2017) or that product availability mediates sustainable purchase decisions (Nguyen et al., 2019). Although this does not provide direct evidence for a payoff bias (Schlag, 1998), it suggests that participants are factoring payoffs into their decision making. Also, participants reported respecting those who behaved sustainably, which may further motivate those influenced by personal payoffs.

There was also indirect evidence an influence of conformity (Henrich & Boyd, 1998). The second highest rated barrier was "*Others around me are not behaving sustainably*", suggesting participants were sensitive to the sustainability of others around them. Participants also generally agreed that they would join in with the sustainable behaviour of others around them. This suggests that conformity could increase or decrease sustainability, dependent on whether sustainable behaviour was the majority or minority behaviour. Social norm studies have identified similar patterns. Injunctive norms (of what do most people approve) supporting sustainability generally encourage sustainable behaviour (Goldstein et al., 2008; Thøgersen, 2008) while descriptive norms (what do most people do) can also be effective (Andor & Fels, 2018), they occasionally produce boomerang effects where individuals decrease their sustainability towards the norm (Schultz et al., 2007).

We found conflicting evidence regarding the influence of friends and family on participant's sustainability. On the questionnaire, participants rated friends and family as likely influences on their sustainability, in addition to perceiving them to hold stronger environmental norms than the public. Similarly, in Slovenia, a mediator model showed the descriptive and injunctive norms of friends and family predicted personal sustainable behaviour and mediated the relationship with country level sustainability norms (Culiberg & Elgaaied-Gambier, 2016). The influence of friends and family may imply the use of a familiarity bias (L. A. Wood et al., 2013) or a similarity bias (Mahajan & Wynn, 2012). As this data is purely descriptive, it is not possible to distinguish these two possibilities. However, the perceived environmental injunctive norms of friends, family, or Durham students did not predict choosing the sustainable voucher. This may reflect an inconsistency between participant's perceptions and the reality of the influence of others (Nolan et al., 2008) or that the norms measured by the questionnaire were not relevant (or salient) for participant's voucher choice.

The pattern of results from the behavioural measure and vignettes was extremely unexpected. The vignettes were not just ineffective but detrimental to sustainability. All three vignettes reduced the probability that participants opted for the sustainable voucher compared to the control condition. While it is not clear why this was the case, here we describe some possibilities. One possibility is demand characteristics, where participants could have tried to behave contrary to what they perceived to be the expected behaviour. While this is possible, experimental evidence suggests that participants generally try to help the experiment when they are aware of the research goals (Nichols & Maner, 2008). Alternatively, the vignettes may have discouraged participant's sustainable behaviour if they felt the behaviour portrayed was unachievable or may have encouraged them to free ride after reading about the efforts of others. This conclusion is also tentative given social prompts have been effective in other sustainability settings (Bergquist & Nilsson, 2016; Goldstein et al., 2008; Player et al., 2018).

A further additional concern relates to the control condition used in this study. In the control condition, participants did not read an unrelated, but equal length vignette and instead received no vignette. Although a different study employing a similar method found no difference in participant's responses between different control conditions (van de Vyver & Leite, 2022) we cannot discount the possibility that this may have affected our results. For example, it may have induced a feeling of increased time pressure and encouraged participants to choose the more familiar voucher (ASOS). As such, some caution should be employed in the comparison of the vignette conditions with the control condition.

Despite the null effect of the vignettes, the pattern of results between experimental conditions warrants further discussion. The lowest probability of selecting the sustainable voucher was in the payoff condition. Given that sustainable behaviour is more costly than less sustainable behaviour, the vignette may have primed participants to evaluate their decision in terms of relative payoffs, as the clothing from Rapanui is more expensive, on average, than those from ASOS. This is consistent with other experiments, where a payoff bias has been shown to exert a negative effect on levels of cooperation in experimental settings (Burton-Chellew, El Mouden, et al., 2017; Burton-Chellew & West, 2021). Both the prestige and conformity framed vignettes reduced the probability of selecting the sustainable voucher relative to the control by a similar degree, though the smallest drop was associated with the prestige condition. While this could suggest that prestige framing was more effective than conformity or payoff framing to encourage sustainable behaviour, when considered alongside participant's self-reported sources of influence, it seems more likely that this further supports the relatively small influence of prestige on sustainable behaviour.

Besides the influence of others, there were several other trends in participant's responses. Participants rated "*I feel that someone else will take care of it"* and *"I do not know how to behave more sustainable"* among their three lowest barriers. This suggests an awareness of how to be sustainable and that they felt a personal responsibility to do so. This contrasts with other studies which find knowledge as a barrier to sustainable behaviour (Axon, 2017; Young et al., 2010), but is consistent with student populations that show good knowledge of and positive attitudes towards sustainability (Emanuel & Adams, 2011; Levine & Strube, 2012). Participants also perceived stronger environmental norms within Durham University students than the general public. While the patterns associated with age are mixed (Tripathi & Singh, 2016), education (specifically, University education) is often positively related to environmental behaviour or commitment (Cotton & Alcock, 2013;

Tripathi & Singh, 2016). Our results are consistent with this picture, as participants held positive environmental norms and perceived strong environmental norms within fellow students.

Our Bayesian model found good evidence for the influence of 5 of the 9 considered predictors on selecting the sustainable voucher. Perceived environmental concern, acting sustainably and being vegan or vegetarian were positively related to selecting the sustainable voucher. Identifying as female and political conservatism were negatively related to selecting the sustainable voucher. Perceiving barriers and perceived environmental norms of family were also weakly negatively related to selecting the sustainable voucher but were more uncertain. Diets which substantially reduce the amount of animal products generally reduce greenhouse gas emissions over traditional diets (Chai et al., 2019) which is one motivation for adopting a vegan diet (Ghaffari et al., 2021). Spill over between these domains could account for the pattern of results related to diet choice seen here (Tiefenbeck et al., 2013). The negative relationship between identifying as female and selecting the sustainable voucher was somewhat unexpected, as females typically (though inconsistently) show higher environmental concern and environmentally friendly behaviour (Tripathi & Singh, 2016). Moreover, in this sample, females generally had higher PCA components scores than non-females on acting sustainably and perceived duty and lower PCA scores on perceived barriers (SI 3). One explanation for the apparent unsustainability we observed in females might be that females had greater familiarity with ASOS than non-females and consequently were more likely to select it. Alternatively, this may be explained by the relatively few non-females present in our sample, as those that took part may be more environmentally friendly than is typical of the population. The pattern with political conservatism supports other studies, which find political conservatives are less environmentally conscious than political liberals (Cruz, 2017) and less responsive to energy interventions (Costa & Kahn, 2013).

Study one allowed an exploration of our research questions in a structured way, in addition to the behavioural measure and experimental manipulation. Using semi-structured interviews in study two allows us to explore the themes and patterns of results identified in study one in greater detail.

6.6. Study 2 Methods

Study two employed semi-structed interviews to explore participant's attitudes and perceptions towards sustainable behaviour in more detail than was possible using a questionnaire. The sample was comprised of individuals who completed the questionnaire and indicated their interest in a follow up interview. Because of its flexibility, this study employed thematic analysis (Braun & Clarke, 2006, 2019). The specific kind of thematic analysis used is described below.

Epistemologically, we adopted a constructionist approach which assumes that participant's accounts reflected their subjective interpretation of reality (Scotland, 2012). This was appropriate, as participant's interpretations of sustainability were of specific interest for our research questions. We combined an inductive and deductive approach when coding and generating themes. Specifically, exploring participant's attitudes and perceptions towards sustainability (RQ2) was better suited to an inductive approach whereby codes and themes were developed without preconceptions about potential themes. Whereas exploring the influence of others (RQ1) was better suited to a combined inductive/deductive approach. This was because RQ2 was generally broad in focus while RQ1 specifically sought to assess the evidence for payoff bias, prestige bias and conformity. However, to not disregard other valuable insights offered by participants, themes were also generated inductively.

6.6.1. Participants

Seven participants took part in the interviews and were recruited via the email addresses they provided in the questionnaire. Though there were no exclusion criteria, many participants declined to take part when contacted, hence the sample is somewhat self-selected. Table 6.7 shows the demographic information collected in the questionnaire (all names given are pseudonyms). There were 6 Females and 1 Male (median age 21 years), and they were all current Durham University students from a variety of departments. Interviews were carried out on a one-to-one basis via Zoom. Two of the participants (Zoe, Diana) knew who I was before taking part and I had previously spoken with Cassie in an unrelated context. Importantly, these participants had no additional information versus the other participants regarding the nature of the research.

Table 6.8. *Participant demographics (names provided are pseudonyms).*

Name	**Gender**	**Age**	**Department**	**Year of study**
Zoe	Female	19	Anthropology	First
Caitlyn	Female	21	Geography	Third
Ed	Male	21	Sociology	Third
Annie	Female	20	Philosophy	Second
Evelyn	Female	22	Sociology	Third
Cassie	Female	30	Education	Fourth +
Diana	Female	20	Anthropology	First

6.6.2. Materials

A semi-structured interview guide (SI 3) was used during the interviews which was based upon trends identified within previous studies. For example, questions prompted discussions about participants understanding of and attitudes towards sustainability, what barriers they feel to their sustainability and the extent that other people influence their behaviour. Being only a guide, interviews frequently deviated from the order the questions were presented in and additional follow up questions were put to participants following their responses.

6.6.3. Procedure

All interviews took place online (Zoom). The audio from the interviews was recorded using Open Broadcaster Software and Zoom's built-in call recorder. Participants were sent the information sheet and consent form (SI 3) to read ahead of the interview.

Upon arrival into the Zoom meeting, several minutes of general chat was encouraged to build rapport. Then, the interview process was explained, and their informed consent was obtained when the recording began. Participants were advised that they could abstain from answering any question (though, none did) and that they would be referred to by a pseudonym in subsequent reports. Once all topics had been covered and the conversation had reached a natural end, the recording was stopped, and participants were sent a debrief form (SI 3). Interview lengths varied from 34 to 59 minutes (average length 46 minutes).

6.6.4. Data Analysis

Following data collection, the interviews were transcribed by RW. Initial transcription was done by ai (Otter.ai, 2021). Errors were then corrected, and punctuation was removed by listening to the recording. The data was transcribed in an intelligent verbatim way which faithfully recorded what participants had said but with utterances (erm, um etc.) and obviously erroneously repeated words removed. Pauses or cases where participants trailed off were retained as this often helped understanding. This approach was most appropriate for our research questions as the interest was exclusively in the content of the interviews (Lapadat, 2000).

Following familiarisation with the data during transcription (during which, initial thoughts were noted) the data were coded using NVivo (March 2020 version). Initial themes were developed and refined in an iterative way throughout the data coding process, such that existing themes were continuously refined, expanded or combined as new themes were developed. Coding and themes were developed primarily at the semantic level to record the ideas expressed directly by the participants, though some codes did consider deeper latent concepts and ideas (Braun & Clarke, 2013). This choice reflected the primary goal of exploring participant's perceptions of sustainability as they understood it. Codes and subsequent themes were mostly generated inductively, though codes pertaining to the specific social learning strategies (conformity, prestige and payoff) were coded deductively. Anonymised transcribed interviews are available here (https://osf.io/frpnu/?view_only=4bfbaccdf7d142fa87093f5aa5bb4ad8).

6.7. Study 2 Results

The coding process generated 4 themes, split between social and non-social influences on sustainable decision making (Figure 6.3). Despite this separation, many themes were closely related and elements of one theme often overlapped with another. Generally, it was apparent that sustainability is a broad and multifaceted topic where every participant offered unique (but overlapping) insights and priorities. All quotations provided are referenced by pseudonyms.

The two themes relating to the non-social influences on sustainable behaviour are "Mechanics of sustainability" and "Sustainability as a personal duty". Mechanics of sustainability was split into the sub themes of "Complexity of sustainability" and "Barriers to sustainable behaviour", related to the challenges participants identified (both for themselves and others) surrounding sustainable behaviour and the variety of other considerations and issues that exist alongside sustainable behaviour. "Sustainability as a personal duty" reflected the largely consistent intrinsically held belief that sustainable behaviour is important (for themselves and others) and an expected social norm.

The two further themes were generated in relation to the social influences on sustainable behaviour were "Perceiving social information" and "Using social information". Perceiving social information was split into the sub themes: "Amplification by social media" which reflected how social media platforms served as catalysts for gathering information, shifting attitudes or a means to ostracise those failing to be behave sustainably and "Scepticism about sustainability" which related to the lack of trust towards the concept of sustainability and particularly advertisements or those that advocate for it. The second theme "Using social information" covered ways in which the behaviour and attitudes of others was used to influence participants own behaviours or attitudes. This section concludes with the deductive part of the analysis which investigated the evidence for the use of prestige bias, conformity or payoff bias and how participants understood the concept of prestige.

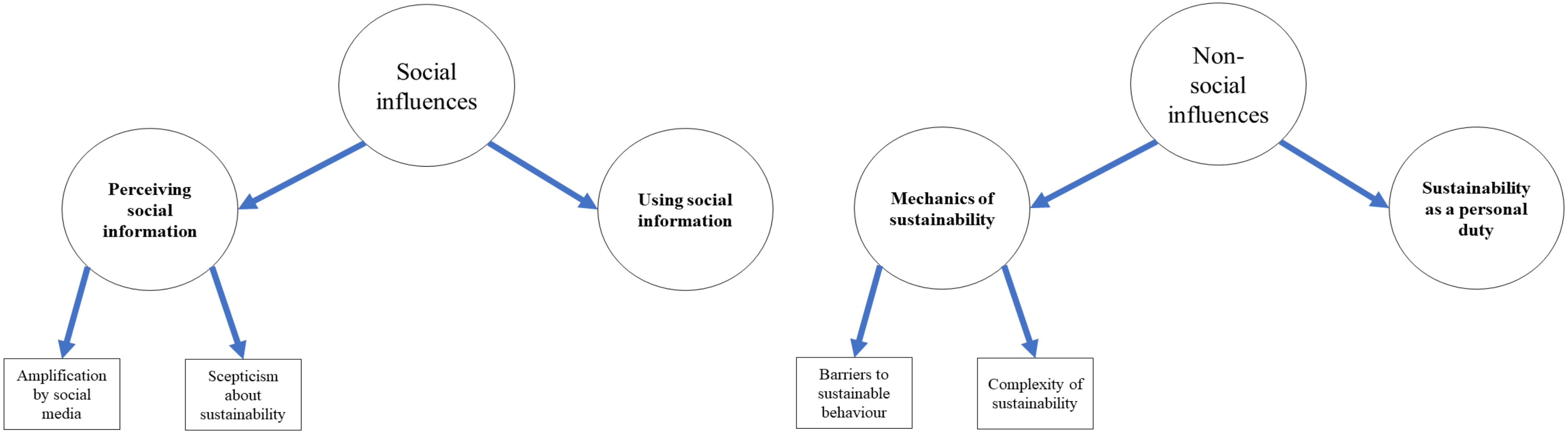


Figure 6.3. *Thematic map showing themes generated during analysis.*

6.7.1. Theme 1 - Mechanics of sustainability: Complexity of sustainability

Each participant's definition of sustainability varied, but they generally centred around behaviours which are good for the environment or that protected future generations. Additionally, participants also indicated that there are different ways one can be sustainable:

"Well I'm not an expert in environment but my understanding is well being sustainable would be being good to the environment and basically it means that maybe we helped to protect the environments you saving the energies you recycling and all those good things to help with the environment (Cassie)"

Caitlyn felt that sustainability was not just about the environment; she preferred a looser definition of "*caring more for the future and the present (Caitlyn)"* which she explained in the context of school children during COVID-19 lockdowns. Sustainability was also understood as part of a wide-reaching tapestry of other societal issues. Diana highlighted a shift in societal understanding from only considering the impact on animals to *"capitalism and actually human behaviour (Diana)"* and Ed described sustainability as a *"long term macroeconomics problem (Ed)"*. Discussions also touched upon other issues such as food waste, pollution, poverty and economic development. Evelyn described how sustainability is often at odds with people's other desires:

"Well like I said before doing my best because like I really wanted this particular brand of chocolate and you know I'm doing my best I want chocolate and I want to be sustainable you know (Evelyn)"

Besides protecting the environment, participants reported many other motivations for behaving sustainably including saving money, improving health, feeling good and social pressure.

6.7.2. Theme 1 – Mechanics of Sustainability: Barriers to sustainable behaviour

Participants described several barriers that they perceived for themselves or for others and alluded to other factors more implicitly. Barriers mostly fell into one of two categories: lack of knowledge or lack of ability. Lack of knowledge was a common concern as it made it difficult to know how best to improve one's sustainable behaviour:

"there's probably tonnes of things that I could be more sustainable with but I just haven't stumbled across the information telling me how to do that yet (Annie)"

For Caitlyn, she felt a lack of understanding of the eventual benefit of her behaviour discouraged her from doing it. She went on to explain, knowledge alone was not enough: *"even though I have the knowledge I'm not really acting on it (Caitlyn)"*. Diana also reflected on this for her family: *"I think it's a lot of talk and no action necessarily yet (.) but I mean they've always been aware of it [sustainability] (Diana)"*.

The effort required to be sustainable was another frequently cited barrier by participants, which Caitlyn linked closely with a scepticism about how her behaviour had any benefit:

"whereas if it's becoming effort on my part especially when I can't see the direct influence of my effort then it's a bit annoying thinking like I've just wasted time doing this and it probably doesn't matter anyway (Caitlyn)"

Evelyn spoke about the difficulty of always recycling items properly if she was tired. But she also believed that making sustainability generally more convenient might be enough to persuade most people to do it:

Evelyn

"but when you're provided with that infrastructure it kind of just integrate it into your everyday behaviour and it doesn't... It takes a lot of burden off everyone to keep at it"

Interviewer

"Yeah so you think I guess given the convenience you think most people probably would go along with it at least to a point?"

Evelyn

"Yeah"

However, changes in behaviour always risks meeting resistance if some people are unwilling to change. Ed suggested this might hold back information campaigns which "[*depend] on one's willingness to be more sustainable or the time or the efforts that they're willing to spend (Ed)".* Diana recalled initial negative perceptions towards veganism: *"I remember a while ago when kind of veganism first came in people were very dismissive of it and I was as well (Diana)".* Zoe also commented that because of her lack of interest in cars and driving she did not *"think about the sustainability of what I'm doing (Zoe)".*

Finally, it may not always be possible to change your level of sustainable behaviour. Being an international student, travelling to and from her home country was not something Evelyn could *"realistically help (Evelyn)"* and Annie described being unable to afford to purchase more sustainable swimwear. Zoe believed that some degree of unsustainability was essentially unavoidable:

"I don't think you could really (.) survive in this world or the world you know the kind of society today without doing some kind of behaviour that's going to be not very sustainable (Zoe)"

This has been exacerbated by COVID-19, which required businesses to switch to disposable packaging amidst public health concerns. In Durham University, college meals were being offered in this way. Several participants noted this and although they agreed it was necessary, regretted the additional waste this produced.

6.7.3. Theme 2 – Sustainability as a personal duty

All seven participants in some form or another expressed the belief that sustainability was a desirable goal even if an individual's efforts may be small:

"it's more about like keeping at it rather than doing it perfect every day and also it's over a longer period of time so it gives a longer time period for you to build up the effect so yeah ((pause)) but the important thing is to do it no matter how small it is (Evelyn)"

Cassie had noticed advertisements presenting sustainability as the morally right thing to do, which in turn affected how she felt about sustainability and Annie felt guilty or sad if she noticed herself failing to consider the impact of her actions. Even for Caitlyn who said: *"I don't care enough about it [sustainability], which is bad (Caitlyn)",* went on to say that "*I would expect more from myself to be honest (Caitlyn)"* which implied that she felt she ought to be more sustainable than she was. A perception of this responsibility also seemed to apply to others. Ed believed that *"more and more people care about sustainability (Ed)"* and that sustainability *"is an essential part in their life (Ed)"* which Caitlyn described as: *"a sort of bar that's already set you have to be sustainable to a certain extent (Caitlyn)"*. Being sustainable can bring positive feelings, which for Evelyn was one of her primary motivations. The expectation to behave sustainably of a held belief that sustainability is a moral duty, where success brings satisfaction and failure, guilt.

Despite this, participants acknowledged that perfection within sustainability is rarely possible. Ed understood this meant accepting that he might not be able to act upon all the information he sees online and for Zoe that she needs to be constantly looking for ways to improve her sustainable behaviour. Above all, participants stressed the importance of being realistic in their expectations and instead favoured an attitude of doing what you can. Diana summed this up using an article she remembered reading from a family friend:

"I read an article the other day written by a family friend and he was just kinda saying you know if you say your veggie but you know you go and get some McDonald's chicken nuggets when you're

drunk on a night out or you know you try and (.) you know (.) think environmentally consciously but you still want to go on holiday because you've never been to India or something like that you know everyone's trying their best you know we can't all be perfect so you know just do your best and don't really... kind of ignore what the critics might say or don't be too self-critical on yourself because you know you can't necessarily avoid certain things (Diana)"

6.7.4. Theme 3 – Perceiving social info

Before discussing the specific themes in participant's responses, it is useful to highlight how participants perceived the attitudes of others. Overall, participants believed others care about sustainability and that is has become more common, especially in the last decade. Both in terms of general attitudes and specific behaviours such as Veganism and avoiding plastic bags:

"over the past 10 years I think it's [sustainability] really shot up and especially in these past couple years it's really... people have... it's definitely like more widely spoken about (Diana)"

Zoe also echoed what other participants expressed about themselves and suggested that people internalise the guilt associated with poor sustainability:

"I do think most people in society these days at least feel slight bit of guilt about kind of sustainable behaviours like there is more of a pressure these days to act sustainably and you will be kind of held accountable for it up to a certain degree (Zoe)"

There remains variability in terms of how much people care, which Cassie noted within her group of friends and Evelyn within her family. Participants offered a variety of explanations for this: some suggested generational differences (specifically, older and younger teens are less sustainable) while others suggested socioeconomic background or poor education. Annie believed that those brought up in households that are less sustainable were prone to continue this behaviour into adulthood, which she stressed deserved greater attention:

"I don't think people grow into like these adults that just don't care about anything if they've had like been taught about it properly as kids a lot of the time (.) so to me it's more of a reminder that that side of things needs so much more focus (Annie)"

6.7.5. Theme 3 – Perceiving social information: Amplification by social media

Social media was one of the primary ways participants engaged with information about sustainability *"I get most of it [information] from a screen basically and from either social media or on a computer (Diana)".* This had benefits, as it granted participants access to a wealth of information unavailable otherwise. Ed for example follows a particular organisation in Durham which posted a range of valuable information around sustainable strategies. More generally, participants thought that sustainability was more common than unsustainability on social media, which Zoe suggested may establish an "*all-encompassing pressure (Zoe)*" on others through viral trends.

Social media also offers influencers and activists a platform to spread their messaging which was the primary way participants engaged with these individuals. Conversely, online platforms can become an arena for individuals to criticise others for failing to be sustainable. This was something that Caitlyn observed often on the platforms she used:

"online it's very easy to jump on a bandwagon of like getting angry at someone for not doing something (.) yeah so I think like social media and stuff has probably contributed to that (Caitlyn)"

Zoe had also observed this on TickTock where in some cases following online arguments users had *"deleted their video or that even their account if it was really drastic (Zoe)"*. She speculated that this sort of shaming *"has quite a bit effect (Zoe)"* but did also say *"they could definitely get just angrier (Zoe)"*. Whether the effect of these interactions was positive or negative therefore was not clear.

6.7.6. Theme 3 – Perceiving social information: Scepticism about sustainability

Several participants expressed a general lack of trust in the sustainability they observed. Diana described sustainability as being "*greenwashed (Diana)"* or *"chucked about here and there as a sort of buzzword (Diana)"* which she explained in advertising where it is common for products to claim they are sustainable even though many aspects of its production are highly unsustainable. Caitlyn was concerned that sustainable products must come with some sort of catch. This links closely with social media, where sustainability could be pushed as a trend or fad. Evelyn used the term *"performative sustainability (Evelyn)"* to describe cases where people (often celebrities) make posts implying they behave more sustainably than they do for likes online. Ed also noticed this during the Earth hour event and was similarly unconvinced:

"some celebrities are promoting the earth hour to remind people to switch off the lights in the house […] I just think that they promote it on social media as they push a fad to remind their fans or people on social media (Ed)"

Zoe felt this applied to celebrities and politicians more generally, who seek to present themselves as acting sustainably "*because it's a trending and important thing (Zoe)*" despite being "*not one of their main concerns (Zoe)"*. Annie offered a unique viewpoint in this respect as she extended her scepticism to any individual acting exceptionally sustainably:

"I feel like if someone's going above and beyond that's cool that's great but when it's someone's like entire life we [her and her friends] were saying like it almost feels like a red flag ((laughs)) like if someone cares that much it's kinda like you question it a little bit (Annie)"

This raised an interesting question about whether sustainability being trendy, or a fad is a negative thing. Annie went on to say she was *"not entirely sure that authenticity even matters because like if they're having a good impact then surely it's fine (Annie)"*. What was notable was that, aside from advertisers, no participant ever expressed any concern about the trustworthiness of information they received, only the extent they could trust others' sustainable behaviour.

6.7.7. Theme 4 – Using social information

While all participants offered at least some examples of ways their sustainable behaviour had been influenced by social information, the extent this occurred and how this was perceived varied substantially. Table 6.8 shows an interpretive breakdown of how important each participant felt the influence of others were overall and whether this happened consciously. The importance of others was judged based on the answers participants gave to questions asking them to reflect on the factors that influence their decisions and those which explicitly asked about the role of others. Whether this happened consciously was inferred from the examples participants gave and explicitly asked of them during the interview.

Generally, participants suggested poor sustainability was more memorable than good sustainability, but good sustainability had the largest influence. Sustainability was usually not memorable unless whatever was observed was very different. Participants highlighted a range of sources they used to observe sustainable behaviours or acquire information including friends, family, activists and documentaries.

Table 6.9. *Interpretive breakdown of the influence of others on sustainable behaviour of participants.*

Participant	Overall influence of social information	Conscious or unconscious	Supporting quotation
Zoe	Secondary to other factors	Conscious	*"there is a bit of the other people influence but I think also I do make sure I'm on track as much as I can be"*
Caitlyn	Little overall influence	Unconscious	*"I feel like just generally I don't get influenced by other people that much"*
Ed	Overall strength unclear	Conscious	*"I think peoples from what people behave or act affect me affects my motivations or ways to be sustainable"*
Annie	Rated as a large influence on attitudes but not behaviour	Unconscious	*"I don't think it's so much of like a mimicking behaviour as it is like having this general conversation and making you feel more obligated to get involved almost"*
Evelyn	Little overall influence	Unclear / unable to judge	*"No it [other people] doesn't influence me ((pause)) I'm just gonna keep doing what I do"*
Cassie	Secondary influence	Mostly unconscious	*"which out of them [influences] do you think is going to be the strongest determinant on sustainable behaviour?"* *"I think both education and economics (.) they kind of in combination play a role in (.) trying to lead people"*
Diana	Catalyses other factors	Mixture but mainly unconscious	*"Yeah I think it's probably the influence of others [...] others then provides you with the information to then potentially mitigate any inconvenience"*

Participants offered more examples of positive than negative direct influences on their behaviour. A large part of this was gaining knowledge from other people. Ed learned about and was inspired by vegan recipes posted by friends from his church and Zoe used social media for inspiration from her competitors for ways to improve her business's sustainability:

"I have that small business that I have and I follow lots and lots of other accounts I find myself most of the time if I'm looking at what they're doing thinking about whether what they're doing is sustainable or not (.) not judging them mainly looking for ideas (Zoe)"

Cassie also offered an example where she adopted a behaviour she was told about, despite not understanding why it was effective:

"I don't know how it works but I heard that maybe reducing the amount of emails can help the environment as well but this is one part I don't understand but I think there must be a reason so sometimes I follow that as well (Cassie)"

There were some limited examples of how social influence had a negative effect on participants' behaviour. For instance, Cassie used a disposable coffee cup following her friend's preference and Diana consumed more meat according to her brother's diet preference. Annie felt that online trends may incentivise unnecessary clothing purchases:

"I definitely do think like I buy too many clothes a lot of the time like do I really need that am I just following a trend (Annie)"

Besides their behaviour, participant's attitudes were also influenced by others, usually in a positive way. Diana was inspired by her friend to shop in charity shops for new clothes, but Cassie suggested witnessing especially sustainable or unsustainable behaviour in her friends would increase her awareness either way. Despite the scepticism that many participants expressed surrounding sustainable advertising, Evelyn nevertheless found it encouraging to some extent:

"when I'm shopping and I see all these sustainable claims printed on a product I think that influenced me a bit I'm like okay the big corporations are doing this I kind of see hope (Evelyn)"

Participants did however highlight some negative consequences of social information. Participants reported that seeing poor sustainability often brought with it negative emotions such as anger or general hopelessness and Caitlyn suggested this further dissuaded her:

"I guess one action that I take is such a small action and it's got to be something that's collective and sometimes when you see other people not do something you think (.) why should I bother kind of thing (Caitlyn)"

Zoe and Annie also felt similarly in response to witnessing good sustainability which, at times, they felt could remove people's responsibility of behaving sustainably themselves:

"If I see someone's doing something good I'm like okay faith restored in humanity that's fine I can relax for a second (Zoe)"

There were other cases where participant felt that social influence had no effect. Within product advertising, in addition to doubting their trustworthiness, price was a greater factor in purchase decisions than sustainability. A product being sustainable could act as a tiebreaker where participants had no other preference. Campaigns promoting sustainability were described as not very memorable, unless it was very unique:

"I'll maybe be a bit more kind of aware and conscious to it [advertising campaigns] but then it'll go and I'll sort of forget about it (Diana)"

Other's behaviour was also less likely to have an impact compared to systematic barriers in adopting the behaviour. Evelyn suggested that she would be unlikely to copy a different method of recycling unless it was significantly more convenient than what she did already and Cassie suggested that being around friends who were more tolerant might moderate their influence:

"although she wants me to behave in such a way but I would not lose her or I would not make any bad impression on her I would maybe not be influenced (Cassie)"

6.7.8. Social learning strategies

Finally, this section considers the evidence that participants used a payoff, conformity or prestige bias in their sustainable decision making. It should be noted that interview questions explicitly asked participants about high status individuals but not conformity and payoff. There were very few indications of the conscious use of any of the three strategies. As such, the role of any social learning strategies must be inferred from participant's answers, and it is not possible to discount that participants were using other social learning strategies than those considered.

As sustainability is usually costly it is not surprising that participants would not use a payoff bias to adopt sustainable behaviour. However, the frequent mention of effort, convenience, and costs as barriers to sustainability suggests that participants are factoring payoffs into their decision making. For example, Cassie opted for a sustainable washing tablet due to financial savings. Additionally, several participants believed prizes or economic incentives would incentivise others.

Participants offered comparatively more indication of a sensitivity to conformity information. In most cases, the attitudes of participants' friends closely aligned with their own. Particularly for Annie, this was the biggest reason for her sustainability. This was accompanied by the general perception that most people feel sustainability is important and are engaged with sustainability to some degree. This may have influenced participants in an unconscious way. Unlike payoff bias however, participants did offer some limited examples which implied a more direct influence of conformity. After moving from her home country, Cassie shifted her internal norm towards the substantially different UK norm and Evelyn adopted a different recycling technique after observing her neighbours and mother. To promote sustainability generally, Zoe suggested that the repetition of sustainable messaging and instructions should be a key element.

A clearer picture was available for participants' sensitivity to prestige (or status). Only Cassie suggested that high status individuals influenced her, but even she could not recall any examples. Other participants either suggested that they had no impact, or that their impact was minimal. In general, this could mostly be explained by a lack of trust in their credibility or that their behaviour is not relatable for regular people:

"we live so differently that means we eat differently we travel differently we do things differently of course we're gonna like do sustainable things differently (Evelyn)"

This became more apparent when participants described the individuals who did have some tangible impact. David Attenborough and Greta Thunberg were both frequently cited sources of inspiration and information, precisely because they were perceived as genuine:

"I think sustainability has become sort of a slightly fashionable sort of thing now whereas he's [David Attenborough] sort of always been fighting for it (Diana)"

Otherwise, it was clear that information from high status individuals was not rated with any additional weight than information from other sources. Annie went as far as saying she would rate information from a *"random person on the street (Annie)"* equivalently to information from a high-status individual. The exception was Cassie who suggested she would value information from high status individuals because she would "*trust their knowledge and their skills (Cassie)"*.

Participants' definitions of how status differed to dominance varied substantially. All participants agreed they were independent but related constructs, though this was where agreement ended. Some participants suggested that status was a perception granted by others and dominance was asserted by the individual while others suggested the opposite of this. Likewise, some participants suggested dominance undermined trust where others suggested dominance enhanced it. There was also little agreement on which would have the greatest effect on sustainability:

"the more seemingly like you know ((pause)) totalitariany you know way […] maybe people are less swayed by that person but they're more feel like they have to listen or do what they're saying whereas the high status you're more likely to listen and understand and agree and stuff like that because they seem to know what they're talking about (Zoe)"

"I think if someone's able to assert dominance without status or having a position of status you know it's sort of a bit more meaningful and they're either passionate about it or passionate that then you do something to change (Diana)"

6.9. Study 2 Discussion

Using semi structured interviews, we investigated the extent that Durham University students used social information to inform their sustainable behaviour (RQ1) and their general perceptions towards sustainability (RQ2). Using thematic analysis (Braun & Clarke, 2006, 2019) we identified four themes. Two related to non-social aspects of sustainability and two related to the social aspects of sustainability. Theme one, "Mechanics of sustainability", reflected the barriers participants faced in their sustainability and the other issues and considerations participants raised around sustainability. Theme two, "Sustainability as a personal duty", reflected the expressed belief that sustainability was a moral obligation that everybody has a responsibility towards. Theme three, "Perceiving social information", captured how participants viewed (and were sceptical towards) sustainability in others and how this was affected by social media. Theme four, "Using social information", referred to the ways in which participants used social information to inform their attitudes and behaviour and by whom they were influenced. Consistent with study one we found some indirect evidence for payoff bias and conformity, but little indication for any influence of prestige (RQ1). More generally, participants held positive injunctive environmental norms and felt a strong obligation to behave sustainably (RQ2).

With respect to social influence (RQ1), overall, participants provided more examples of cases where the behaviour of others had no effect on their behaviour than examples where it did and generally did not believe that the behaviour or attitudes of others had much impact on their behaviour. Within sustainability, it is common for individuals to underestimate the influence others have on their sustainable decision making (Frey & Meier, 2004a; Nolan et al., 2008). This is in spite of the effectiveness of many normative campaigns within sustainability (Andor & Fels, 2018) and other applied contexts (Yamin et al., 2019). Participants additionally suggested that the influence that occurred, took place unconsciously. This links with wider debates within the cultural evolution (Heyes, 2016; R. Kendal et al., 2018) and social influence (Nolan et al., 2008) literature more broadly as to what extent social learning occurs consciously. However, there are several reasons why purposeful social learning could be less useful in a sustainability context, which may explain the lack of social influence described by participants. Because people's precise circumstances vary (in terms of how someone can improve their sustainable behaviour), socially acquired information may be less useful, which in turn favours asocial over social learning (Feldman et al., 1996). Alternatively, psychological factors (such as social norms) are argued to be weaker predictors of sustainable behaviour when sustainable behaviour is less constrained by systematic barriers (Stern, 2011). This would imply that systematic barriers have a bigger impact on individuals' degree of sustainability than social influence.

Of the social influence participants discussed, social media was their primary source of information. Social media has been shown to promote sustainability through influencers spreading information and active learning and participation from users (Chwialkowska, 2019; Shrivastava et al., 2021). In young people, social media use has been linked to greater green purchase intention (Bedard & Tolmie, 2018). Social media has also simplified the buying and selling of second-hand goods on sites such as Facebook or Gumtree. Other sustainability trends such as #global citizen (Merle et al., 2019) and freecycling groups (Norbutas & Corten, 2018) have also spread on social media. Though, some participants did note that social media can become an arena for arguments, which also has been shown in online sustainability communities (Andersson & Öhman, 2017).

Most other influences participants described came from friends or family. Previous experiments have shown greater influences from known peers than random individuals on opting for a sustainable prize (Salazar et al., 2013) and normative in-group relevant signs prompted greater decrease in car idling than non-in-group relevant signs (Player et al., 2018). In terms of social learning strategies, this may imply a familiarity bias (Corriveau & Harris, 2009; L. A. Wood et al., 2013). Participants also described cases of negative influences, where witnessing others' sustainability reduced their feeling of responsibility and witnessing others not being sustainable reduced the feeling that their behaviour was worthwhile. This is somewhat similar to patterns observed in normative interventions, where low use households sometimes increase their usage towards higher usage households (Schultz et al., 2007) or field experiments which find antisocial behaviour is more common in environments where there are other norm violations (Keizer et al., 2008; Keuschnigg & Wolbring, 2015).

We now consider the evidence for payoff bias, prestige bias and conformity (RQ1). Participants made no direct mention of using a payoff bias, which was expected because sustainability is usually costly (for example time or money) compared to less sustainable alternatives. However, it may be inferred that participants were factoring payoffs into their decision making based on their mention of time and financial costs as barriers. One participant also reported that behaving sustainably made her feel good, which may provide a payoff motivated incentive to engage in sustainable behaviour. There was a comparatively greater indication for conformity. Many participants reported assorting with likeminded friends and generally felt that others shared their attitude. Additionally, while no participant suggested this explicitly, it is plausible that the duty towards sustainability expressed by participants is social in origin and based on the attitudes of others, given that social norms have been shown to exert strong influences on sustainable behaviour (Allcott, 2011; Andor & Fels, 2018).

The patterns with prestige bias were clearer. Excluding some specific examples, participants reported little interest in high status individuals that endorsed sustainable behaviour. Prestige has been shown to sustain altruistic behaviour theoretically (Henrich et al., 2015) and environmentally friendly behaviour at the leadership level predicts intention to act pro environmentally of staff within a workplace (Blok et al., 2015). Therefore, our results are surprising. Primarily, participants expressed a lack of trust or a feeling that prestigious individuals are not relatable or relevant. Being relatable and consistent has been shown to predict the extent that social media influencers can promote sustainable behaviour (Chwialkowska, 2019). The exceptions were Greta Thunberg and David Attenborough, who participants did suggest were both trustworthy and somewhat influential. This implies participants favoured specific knowledgeable individuals over employing a generalised prestige bias.

The variation participants displayed in the ways and extent to which they used social learning is also notable. While some participants suggested it played almost no role at all in their decision making, others offered several different examples of the way it was used. This pattern is consistent with the emerging understanding within social learning research of the considerable individual differences in the extent that individuals use social learning and the specific strategies (or mix of strategies) that individuals employ (R. Kendal et al., 2018; Rawlings et al., 2017). What was also clear is that individuals did not use social learning indiscriminately. They instead carefully selected which individuals they sought information from, which is consistent with predictions from evolutionary theory (Boyd & Richerson, 1995).

Considering participant's general perceptions of sustainability (RQ2), participants provided generally similar definitions of sustainability centred around protecting the environment and securing future generations. Many participants also stressed that sustainability existed alongside, and was further influenced by, other societal issues such as consumerism, income inequality and education (theme 1, subtheme 2). This is dissimilar to other studies, which found relatively few participants considered the wider social impacts of their product choices (Eberhart & Naderer, 2017) or definitions of sustainability more broadly (Simpson & Radford, 2012). This may be explained by our student sample, as awareness of sustainability is often positively associated with educational attainment (Casey & Scott, 2006; Tripathi & Singh, 2016) and students generally display highly positive environmental attitudes (Yılmaz & Erkal, 2017).

Participants also discussed several barriers to their sustainability (theme 1, subtheme 2). Lack of knowledge, both in terms of how to be sustainable or the benefit of their behaviour, as a barrier to their sustainability was frequently discussed. This is consistent with other studies which also show awareness or knowledge as barriers to sustainable behaviour (Acharya, 2019; Axon, 2017; DEFRA, 2018; Young et al., 2010) or the purchasing of green products (Joshi & Rahman, 2015). However, information-based interventions to increase sustainability show only mixed success (Delmas et al., 2013). It is interesting therefore, that participants reported it so frequently, which may suggest it is necessary for sustainable behaviour but not sufficient. Other barriers included systematic factors such as infrastructure, price or effort/inconvenience, which have also been identified as barriers to purchasing green products (Joshi & Rahman, 2015). Within recycling interventions, making recycling easier was more effective than normative prompts (Carlson, 2001). Many participants rated systematic barriers as their main personal barrier. Participants were also highly sceptical towards sustainability in some contexts (theme 3, subtheme 2), primarily with respect to advertising. Previous studies have shown that distrust of environmental campaigns or the validity of the climate crisis is negatively related to sustainable behaviour (Malodia & Bhatt, 2019). Credibility also mediated the relationship between attitudes and purchase intention (Chen & Chang, 2012) and the influence of celebrity endorsement of sustainable products (Cuomo et al., 2019). This has been compounded by the phenomena of "greenwashing" (where products or companies are misleadingly marketed as environmentally friendly) (Nguyen et al., 2019) which one participant quoted explicitly. For this reason, participants felt advertising was of little consequence to their sustainable practices.

Participants also felt a strong moral duty to behave sustainably and believed it was important to do what you could no matter how small (theme 2). This implied sustainability was understood as a collective action problem which supports theoretical descriptions that formulate sustainability as a prisoner's dilemma (Keohane & Victor, 2016). Participants believed other individuals, for the most part, shared this attitude. This resonates with UK consumer trends which have shown continued expansion of the sustainable consumer market since 2010 (Ethical Consumer, 2020). Despite this, there remains a widely appreciated "attitude-behaviour gap" where people commonly fail to translate sustainable attitudes into behaviour which indicates some continued shortfall in sustainability (Nguyen et al., 2019). Indeed, several participants noted this, claiming that knowledge does not necessarily translate into action.

Overall, this study found little direct support for the use of payoff bias, prestige bias or conformity (RQ1). However, we found indirect evidence that participants considered payoffs in their sustainable decision making and were somewhat sensitive to social norms. Besides this, participants were influenced most strongly by specific or familiar individuals, suggesting a familiarity bias (L. A. Wood et al., 2013). Participants generally held strong injunctive environmental norms and felt it was their obligation to do be as sustainable as they could (RQ2). They additionally cited many barriers including knowledge and systematic factors such as inconvenience or price.

6.10. General discussion

Using a mixed methods approach, we assessed two research questions relating to the sustainable behaviour of Durham University students; the extent that students used social learning or were influenced by social information (RQ1) and their broad overall perceptions of and attitudes towards sustainable behaviour (RQ2). We also evaluated the evidence that participants used (or were influenced by) a payoff bias (Schlag, 1999), prestige bias (Henrich & Gil-White, 2001) or conformity (Henrich & Boyd, 1998). Overall, we found little evidence for the direct influence of any of the three social learning strategies, but some indirect evidence of payoff bias and conformity (RQ1). However, participants in our sample showed a high degree of environmental awareness and generally believed that being sustainable was an important personal obligation or duty (RQ2).

Study 1 showed that other potential social learning biases were more likely to influence participant's behaviour than payoff, prestige or conformity. These included familiarity (Corriveau & Harris, 2009; L. A. Wood et al., 2013) in the form of friends and family and knowledge (R. Kendal et al., 2015; L. A. Wood et al., 2015) in the form of the scientist. Additionally, using an experimental design, participants were not encouraged to be more sustainable when exposed to payoff, prestige or conformity framed information. However, we found indirect evidence of payoff bias and conformity based upon participants rating costs as the most influential barrier to their sustainability and that they perceived more positive environmental attitudes in their friends, family and fellow students than the general public. Study 2 supported this general pattern of results. Participants offered little explicit indication of the use of payoff, prestige or conformity but again spoke broadly about the impact of costs, inconvenience, and the influence of social norms on their sustainability. With regard to prestige, participant's definitions were inconsistent, and they expressed distrust towards individuals of high-status that lacked environmental expertise. Like study 1, participants suggested friends and family were the most likely individuals to influence their behaviour, but the impact of social influence overall was small compared to other factors.

We note some limitations of our sample. Our studies were conducted on students who tend to be more aware of, positive towards, and engaged with sustainability than the general population (Cotton & Alcock, 2013; Emanuel & Adams, 2011; Levine & Strube, 2012). Our sample is further (mostly) restricted to a WEIRD population (Henrich et al., 2010). Our questionnaire sample was substantially gender biased (81% identified as female), which may be relevant as females on average report higher levels of sustainable behaviour than males (Tripathi & Singh, 2016). Although gender was controlled for in our Bayesian model, this may have affected responses to the questionnaire more broadly. Additionally, there were many more students from Psychology and Anthropology (which themselves are generally female biased) than other departments. Independent studies have found that while engineering students generally felt a responsibility towards sustainability (Tang, 2018) this was not the case for business students (Eagle et al., 2015) which does imply the existence of some variation between university departments in terms of attitudes towards sustainability. To our knowledge, no study has systematically compared sustainable behaviour or attitudes between departments, so it is hard to estimate the extent to which this may have influenced our results.

Our interview sample was gender biased (6 females, 1 male) and was comprised of individuals who had already taken the time to fill out a questionnaire about sustainability and agreed to be interviewed. It is possible, therefore, that these participants may have been inclined towards sustainability in a way that other samples may not have been. Therefore, we caution that our results should not be generalised to the overall population. Instead, we favour perspectives offered elsewhere (B. Smith, 2018) which argues generalisability need not be understood in terms of the extent that findings may be extrapolated to other populations. Instead, findings on different populations can be compared to one another to explore how they are similar or different, to produce a more complete picture of the phenomena of interest (B. Smith, 2018).

Also, relevant here is RW's positionality as a researcher. While RW endeavoured to remain neutral and impartial throughout the interview process, as sustainability is an important and topical issue in modern society, it was impossible to completely divorce RWs own opinions from the research. As such the interviews and data interpretation has been conducted through this perception. Furthermore, RWs position as a fellow student may also have impacted the research process. For instance, it may have helped to build rapport and made participants feel more at ease but could have made the interviews less focused. All of this is to say that some caution must be employed in extrapolating these results to other populations. We emphasise the recommendations of Smith (2018) and suggest the themes identified in our study be compared and contrasted with those from other populations.

6.11. Conclusion

Our study aimed to assess the extent that Durham University students used social learning in a sustainability context and their perceptions of sustainability. Overall, we found indirect evidence for payoff bias (based on the cost and inconvenience participants identified as barriers to their behaviour) and conformity (based on the general influence of social norms), but little evidence of prestige bias. Instead, participants were more influenced by familiar or knowledgeable individuals. Further research could address the research questions presented here in different, more generally representative samples to evaluate the extent that the trends identified may apply more broadly. Additionally, focused interventions employing familiar or knowledgeable models may also be effective to spread sustainable behaviour.

Chapter 7: General discussion

7.1. Introduction

In this thesis, I sought to evaluate the impact of social learning on cooperative behaviour. To this end, I employed a mixed methods approach comprised of two online experiments (developed in Dallinger, (2020), an agent-based model, a questionnaire and semi-structured interviews. These methods allowed me to investigate my primary research question within and beyond an experimental context and evaluate the extent that the experimental findings were applicable in a real-life context. With the exception of the spite study (chapter 5) which did not consider prestige bias, each study evaluated the impact of payoff bias (McElreath & Henrich, 2003; Schlag, 1996b), prestige bias (Henrich & Gil-White, 2001; Jiménez & Mesoudi, 2019) and conformity (Henrich & Boyd, 1998; Whiten, 2019a) on cooperative behaviour. Other potential social learning strategies were also assessed as appropriate.

Study 1 (chapter 4) was an online public goods game (PGG) that provided participants access to payoff, prestige and conformity related social information and compared a prisoner's dilemma (PD) with a snowdrift (SD) payoff structure. Model comparison supported a small effect of payoff bias to explain the changes in cooperation observed across rounds with some additional evidence that contributions were higher in the SD than the PD game. Study 2 (chapter 5) focused on spiteful and altruistic behaviour and provided participants with success or majority weighted social information. In this study, I found strong preferences for altruism and no evidence that information source (success or conformity) affected participant's behaviour. Though, there was some evidence that being exposed to altruistic social information encouraged participant's altruistic behaviour. Finally, study 3 (chapter 6) combined a questionnaire and semi-structured interviews to investigate social learning within the applied context of environmental sustainability. Overall, in this study, I found indirect evidence for the use of payoff bias and conformity but little indication for the use of prestige bias. This study also permitted further exploration of other social learning strategies used within a sustainability context.

This General Discussion chapter brings together the general findings of this thesis, focusing on the overarching themes present across the three studies. I discuss the evidence for the use of payoff bias, prestige bias and conformity to inform cooperative behaviour and relate this to the social learning and cooperation literature. I then discuss methodological and theoretical limitations of the work presented here and suggest ways that future research can build upon and develop the findings presented here.

7.2. Social learning Strategies

In this section, I consider the evidence for the use of social learning or the influence of social information more broadly across the three empirical chapters. Sections are split between each of the three primary strategies considered (payoff bias, prestige bias and conformity), followed by the exploratory evidence pointing towards the use of other potential social learning strategies.

7.2.1. Payoff bias

Using a payoff bias (reviewed in section 2.4.2), individuals copy behaviour, or individuals that display a behaviour, with a greater payoff than their current behaviour (McElreath & Henrich, 2003). This is considered critical for cumulative cultural evolution (Mesoudi & Thornton, 2018; Vale et al., 2017).Within social dilemmas that follow a PD payoff structure (Rapoport, 1974), an expected prediction from game theory is that a payoff bias should favour defection and erode cooperation. Evolutionary models where traits are expressed in proportion to their payoffs (Taylor & Jonker, 1978) have shown that this results in full defection in PD games. Experimental PGG studies usually find initially high contributions which decline across rounds of the game (Fehr & Schurtenberger, 2018). This is consistent with other evidence suggesting that providing participants with feedback about how their behaviour was benefiting others caused them to reduce their contributions (Burton-Chellew et al., 2015; Burton-Chellew & West, 2013) and that declines in PGGs were predicted by participants being able to influence their own payoffs (Burton-Chellew & West, 2021). With respect to social learning strategies, three experiments that directly compared frequency with payoff information found that participants were generally most sensitive to payoff weighted information, causing them to reduce their contributions (Burton-Chellew, el Mouden, et al., 2017; Burton-Chellew & Amico, 2021;

Molleman et al., 2014). In sum, when used in cooperative contexts, payoff biases generally result in a decline in cooperative behaviour.

The results obtained in this thesis are consistent with findings identifying the influence of payoff bias on cooperation. Study 1 (chapter 4) which provided access to all three kinds of social information (payoff, prestige and conformity), found the changes in cooperation observed across rounds was most consistent with a payoff bias. Although the overall differences between conditions were small, study 3 (chapter 6), found the payoff vignette had the strongest effect on the probability that participants opted for the sustainable voucher. The interviews, and questionnaire revealed that participants perceived the costs and effort required as a barrier to their sustainable behaviour, providing indirect evidence of a payoff bias. Conversely, a success bias (related to a payoff bias) was no more influential on participants' altruistic/spiteful decision making than conformity (chapter 5). Considering the volume of previous evidence in support of a payoff bias in cooperative contexts, the results from study 1 are not surprising, particularly when a payoff bias is compared to a prestige bias and conformity. Prestige bias and conformity are heuristic strategies that rely on other cues of a trait's (or model's) success when direct payoff information is unavailable (Henrich & Boyd, 1998; Jiménez & Mesoudi, 2019; McElreath et al., 2008; McElreath & Henrich, 2003). As the PGG used in study 1 gave participants free access to accurate payoff information, it was not necessary for them to rely on a heuristic strategy. Conversely, in study 2 (chapter 5), both conformity and success bias were heuristic strategies in the sense that neither could inform participants about the actual payoffs associated with behaving altruistically or spitefully. As such, there was no advantage (in terms of payoffs) in using success over conformity information.

In study 3 (chapter 6), I found that participants considered payoffs within their sustainable decision making. However, this does not qualify as payoff based social learning per se as this was not an example of the copying of traits or individuals based on the payoffs of their behaviour. Instead, the use of payoff information appeared to be related to the general fact that sustainable behaviour, and products, tend to be more costly (in terms of money or time) than less sustainable alternatives. Previous exploratory studies have shown that students (Horhota et al., 2014) and consumers more generally (Joshi & Rahman, 2015) feel price is a barrier to them purchasing sustainable products. Additionally, interventions which made recycling easier were more effective than those using normative prompts (Carlson, 2001) and the adoption of electric vehicles was influenced by logistical factors such as free parking and having travel patterns appropriate for an electric vehicle (Langbroek et al., 2016). The impact of a payoff bias was further supported by the findings from the vignette manipulation, where the payoff vignette had a larger effect on participants decisions to select the sustainable voucher than the conformity or prestige vignette. Overall, the impact of payoff biased social learning found in this thesis is broadly consistent with previous evidence that payoff biased learning is used when individuals have access to accurate payoff information (McElreath & Henrich, 2003).

7.2.2. Conformity

Using conformity (reviewed in 2.4.3), individuals are disproportionately likely to adopt the most common trait within a population (Boyd & Richerson, 1985). Traits are copied such that the frequency of copied traits increases over generations producing characteristic sigmoidal or "S" shaped curves[17] (Acerbi et al., 2016; Henrich & Boyd, 1998). This differentiates conformity within cultural evolution from conformity in social psychology which does not demonstrate disproportionate copying (Bond & Smith, 1996). However, many studies investigating conformity do not demonstrate the disproportionate copying of traits and are referring to a social influence causing individuals to adopt a new behaviour despite their prior contrasting knowledge (Claidière et al., 2014; Grove, 2019; Haun et al., 2012).

The role of conformity on cooperation is mixed and most studies are also restricted to demonstrating social influence rather frequency dependent conformity. Theoretically, conformity can allow populations with established cooperative norms to resist invasion from defectors (Boyd & Richerson,

[17] Though, finding evidence of an "S" shape within a population with respect to a particular trait does not necessarily provide evidence of conformity (Acerbi et al., 2016).

1985). However, laboratory studies have shown that conformity is less influential on cooperative behaviour than reciprocity (Bardsley & Sausgruber, 2005; Romano & Balliet, 2017) and payoff weighted information (Burton-Chellew et al., 2017; Burton-Chellew & Amico, 2021). An applied field study also showed that only some participants' donation behaviours were affected by the frequency of previous participants' donations, suggesting variation in the extent that individuals use conformity (Frey & Meier, 2004b).

Study 1 (chapter 4) showed that conformity information was largely ignored by participants in the PGG. Of the three social learning strategies considered, conformity had the smallest influence on participants cooperation, and the models containing conformity were penalised most heavily by WAIC, indicating poorer predictions out of sample. Study 2 (chapter 5) showed that the source of the social information had little impact on participants' spiteful behaviour. Some indirect evidence of a potential conformity effect was found with respect to sustainable behaviour (chapter 6). Participants reported being sensitive to the sustainability norms of their close friends and, in the case of international students, sensitive to the normative differences between their home countries and the UK. The conformity vignette also decreased the probability that individuals selected the sustainable voucher, to a lesser extent than the payoff vignette. Overall, I found little evidence for the use of conformity.

There are several possibilities that might explain why little evidence for conformity was found in these studies. The first is the relationship between conformity and payoff bias alluded to in the section above (7.2.1.). Using conformity, individuals do not evaluate the payoffs or success of a given trait, only the frequency by which it is expressed (McElreath & Henrich, 2003). Consequently, conformity is less influential when individuals have access to accurate payoff information (Barrett et al., 2017; McElreath et al., 2008; Mesoudi, 2011). This makes sense in the context of cumulative cultural evolution which requires that traits are iterated and improved upon by future generations (Mesoudi & Thornton, 2018; Vale et al., 2017). Strong conformity can hinder the spread of beneficial traits (Kandler & Laland, 2009; Whitehead & Richerson, 2009), and is usually outcompeted by a payoff bias (Grove, 2018) and sometimes random copying (Grove, 2019). Given participants in study 1 (chapter 4) had full knowledge of the payoff structure underlying their behaviour, there was no reason for them to use any other strategy. It follows that the best evidence for conformity was found within sustainability (chapter 6), where the relative payoffs of any given behaviour are less clear (meaning heuristic strategies are more useful). This is similar to the use of prestige biased social learning found in Fiji food taboos, where payoff information was difficult to assess (Henrich & Henrich, 2010).

Under conformity, the most common trait is assumed to also be the most successful (Henrich & Boyd, 1998). In the case of an economic game, it is extremely obvious when this assumption is violated, and the most common trait is extremely unlikely to also be the most successful. In studies 1 and 2, participants could verify this using the payoffs associated with each donation (study 1) or the payoff changes associated with their behaviour (study 2). Although it is true that social learning can sometimes spread maladaptive traits (Mesoudi, 2009), this did not appear to happen in my experiments. Economic games are perhaps not an environment where conformity is useful. Once again, the greater indication of conformity in sustainable behaviour may reflect the fact that it is unclear whether common behaviours are the most beneficial.

Another consideration is that conformity does not necessarily have to act on cooperative behaviour directly to maintain cooperation in a population; conformity could homogenise social norms within a population (either related or unrelated to cooperation) (Boyd & Richerson, 1985; Cialdini et al., 1991), increasing social cohesion and the likelihood of cooperation (Molleman, Quiñones, et al., 2013). This latter possibility was not assessed in this thesis. More directly, conformity could spread injunctive norms (what ought you do) associated with cooperation (Cialdini et al., 1991). Within field research on environmental sustainability, descriptive norms sometimes cause “boomerang” effects where individuals increase or decrease the target behaviour towards the mean (Schultz et al., 2007). Injunctive prompts are generally more effective than descriptive prompts to promote sustainability and avoid the boomerang effect (Goldstein et al., 2008; Player et al., 2018; Schultz et al., 2007). This is consistent with findings from study 3 (chapter 6) which indicated that participants were generally affected by the held norms of others, rather than directly copying their behaviour. Therefore, although in this thesis I found little evidence that conformity directly spread cooperative behaviour, perhaps conformity would cause the spread of cooperative norms.

Finally, the evidence of conformity presented in this thesis did not demonstrate the disproportionate ("S" shaped) copying of traits (Boyd & Richerson, 1985). Generally, it is unclear what to call this weaker kind of conformity. The term "hyper conformity" was suggested to be reserved for conformity in the cultural evolution sense while "linear conformity" referred to the weaker kind (Claidiere & Whiten, 2012). Later work instead calls "cultural evolution conformity" conformist transmission or just conformity, while non "S" shaped conformity is termed Aschian conformity or social influence (Muthukrishna et al., 2016; Whiten, 2019a). Consequently, it is not always obvious exactly what some studies mean by the term "conformity" (Bardsley & Sausgruber, 2005). This is further complicated by my consideration of conformity acting on continuous traits (as the operationalisation of cooperation and spite in studies 1 and 2), as conformity is rarely considered in this way (Morgan & Thompson, 2020). The way conformity was operationalised in this thesis is discussed further in section 7.4.2.

7.2.3. Prestige

A prestige bias (reviewed in 2.4.4) is a heuristic strategy like conformity that is useful when it is not possible to accurately assess the payoffs of different models (Brand et al., 2020; McElreath & Henrich, 2003). As originally defined by Henrich and Gil-White, (2001), generally skilful individuals are granted prestige by followers who pay them deference in exchange for learning opportunities. Consequently, prestigious individuals are influential beyond their domain of expertise (Henrich & Gil-White, 2001; Jiménez & Mesoudi, 2019).

Prestige is theoretically highly beneficial for the evolution of cooperation, particularly when compared to payoff bias and conformity. While a payoff bias results in the decline of cooperation in most circumstances (Taylor & Jonker, 1978) and conformity depends on the initial composition of the population (Carpenter, 2004), prestige suffers neither of these drawbacks: provided the prestigious individual is cooperating and followers are sufficiently likely to copy them, this can promote cooperation in the population (King et al., 2009). In simulations, a larger prestige bias was associated with larger group sizes where cooperation could be sustained and the deference prestigious individuals received incentivised them to continue cooperating (Henrich et al., 2015). Prestige could maintain cooperation under more environmental conditions than conformity (Molleman, Quiñones, et al., 2013). Despite strong theoretical support, there are very few empirical demonstrations of a relationship between prestige and cooperation.

In this thesis I found very little evidence of a prestige bias in relation to cooperative behaviour[18]. In study 1 (chapter 4), prestige information was largely ignored in favour of payoff information. Though, between conformity and prestige, there was greater evidence for prestige. There was also some indication that prestigious individuals were more cooperative, providing some support for theoretical predictions (Henrich et al., 2015). In relation to sustainable behaviour (chapter 6), participants reported not only being insensitive to prestigious individuals, but in many cases were actively dissuaded by them. This was driven by scepticism and a feeling that high status individuals (like celebrities) had little relevance to their own lives. The prestige vignette also had the smallest impact on choosing the sustainable voucher compared with conformity and payoff. As there are very few empirical studies that consider cooperation and prestige, it is more difficult to contextualise these findings within previous work than it is with conformity or payoff bias. However, this thesis did not support experimental work showing status predicted copying of cooperation (Kumru & Vesterlund, 2010) nor did it (unambiguously) support predictions from theoretical models (Henrich et al., 2015; Molleman, Quiñones, et al., 2013). Furthermore, the lack of trust that participants in the interviews expressed towards prestigious individuals (chapter 6, interviews) contrasted with theoretical definitions of prestige, which suggests that prestigious individuals should be trusted based on their status and influential beyond their domain of expertise (Henrich & Gil-White, 2001). Instead,

[18] Note that study 2 (chapter 5) did not consider the role of prestige in spiteful behaviour.

participants from the questionnaire (study 3) indicated that scientists (who possess knowledge and status in a relevant domain) were the most likely to influence their sustainable behaviour.

Like conformity, it is plausible that prestige was not influential due to the availability of payoff information. To avoid repetition, I will not repeat such discussions here (see section 7.2.2). Though it is notable that, unlike conformity, generalised prestige had no effect on participant's views towards sustainable behaviour, despite the lack of access to accurate payoff information. More generally, several trends present in the prestige literature were also present in this thesis.

There is a significant definitional divide between studies in this thesis in terms of how prestige was defined and operationalised. Theoretically, an individual's prestige can be based on their general knowledge or skill or their popularity (Jiménez & Mesoudi, 2019). In study 1 (chapter 4) an individual was considered prestigious if they demonstrated context relevant knowledge on a quiz, while study 3 (chapter 6) focused on popular or high-status individuals (celebrities, social media influencers). This inconsistency is not restricted to this thesis. Investigating the prestige literature reveals a range of ways that prestige has been indicated, including physical looking times (Chudek et al., 2012), virtual looking times (Atkisson et al., 2012), quiz performance (Brand et al., 2020, 2021) or knowledge in an unrelated (Henrich & Broesch, 2011) or related (Jiménez et al., 2018) domain. This suggests ambiguity as to what the term "prestige" means or to what kind of individual it refers to. The exploratory nature of the interviews (chapter 6) is particularly useful here, as participants offered their understanding of what it meant to be prestigious over being dominant. At the most extreme, one participant provided a definition of prestige extremely close to the theoretical definition (Henrich & Gil-White, 2001), while another used an identical definition to describe a dominant individual. This suggests that the concept of "prestige" as it is conceptualised in the literature is not applied consistently within the general population. This could partly account for the divide between the results obtained in this thesis and the theoretical link between prestige and cooperation (Henrich et al., 2015) as not all participants will necessarily perceive the same individuals as prestigious.

Findings associated with prestige are also inconsistent in the literature. In general, while some studies find strong prestige effects, others do not (Jiménez & Mesoudi, 2019). Mostly, this can be explained by the context of the study. For example, when the task did not require expertise (Acerbi & Tehrani, 2018) or the domain of prestige was irrelevant to the observers (Verpooten & Dewitte, 2017) (a concern also shared by my interview participants) this reduced the effect prestige had on participant's behaviour. Alternatively, if the prestigious individual's knowledge is restricted to an unrelated context, this also reduces the probability that individuals use a prestige bias (Brand et al., 2021; Chudek et al., 2012). The real-life cases where influences of prestige have been found seem to be those where prestige hierarchies emerged naturally (Henrich & Broesch, 2011; Offord et al., 2016, 2019). It may be the case that prestige is a powerful force within cultural evolution, but cues must be believable, organic, and relevant for an individual's prestige to affect the probability that those individuals are copied. Nevertheless, in this thesis I did not find good evidence for the use of a prestige bias within cooperative decision making.

7.2.4. Other social learning strategies

Study 1 (chapter 4) and study 2 (chapter 5) were limited to considering conformity, payoff bias and prestige bias, as the social information provided to participants was experimentally manipulated. Study 3 (chapter 6) was unique in this respect, as its exploratory nature meant it could consider additional social learning strategies.

Within an environmental sustainability context (chapter 6), there was more evidence that participants were using other social learning strategies than conformity, payoff bias or prestige bias. In the questionnaire, participants rated the scientist as the most likely source of influence on their sustainable behaviour. It is ambiguous as to what kind of social learning strategy this implies: copying a scientist could represent a prestige bias (Henrich & Gil-White, 2001) or a knowledge bias (R. Kendal et al., 2015; L. A. Wood et al., 2015). While it was not possible to distinguish these two possibilities, based on participants' other responses, it seems more likely that this reflects a knowledge bias. Firstly, the other prestigious options offered to participants ("individual popular on social media" and "celebrities") were rated as the least likely to influence their sustainable behaviour, suggesting participants were not using a generalised prestige bias. Secondly, scientists plausibly hold domain relevant knowledge regarding sustainability, meaning a heuristic strategy like status-related prestige bias would not be necessary (Henrich & Gil-White, 2001). This is supported by the interviews, where participants consistently reported that prestigious individuals were not influential. However, energy audits (which imply a knowledge bias as they involve an expert giving tailored advice to homeowners) are only weakly effective at encouraging housing refits (Murphy, 2014), so it is not clear to what extent my findings may generalise to other contexts.

What was clear, was that participants were influenced by individuals they were close to, or familiar with, such as their friends or family. Again, it is not possible to categorically separate these two possibilities, but this could reflect either a familiarity bias (Corriveau & Harris, 2009; Laland, 2004; L. A. Wood et al., 2013) or a similarity bias (Mahajan & Wynn, 2012) in the sense that a major influence participants cited was from their friends at Durham University with whom they would be demographically similar. This general finding is consistent with several other studies investigating sustainable behaviour. Experiments have found that individuals are more likely to comply with requests if they share some (arbitrary) characteristic with the requester (Burger et al., 2004). Within sustainability, this may explain why compliance with normative requests is higher if they are framed as originating from a familiar group than an external (unfamiliar) group (Goldstein et al., 2008; Player et al., 2018) and why interventions that spread information using a single individual (whom stakeholders can get to know) are more effective than other dissemination strategies (Abrahamse & Steg, 2013). Additionally, information from a trusted source was more likely to promote energy reductions than information from a less trusted source (Craig & McCann, 1978) and celebrity credibility mediated their positive influence on sustainable products (Cuomo et al., 2019).

All three empirical chapters also suggested a further general trend; social information content was more relevant than its source. Study 2 (chapter 5) found that the differences observed between experimental conditions were associated with the content (the behaviour) rather than the source (successful or majority) of the information. In the interviews (chapter 6), some participants noted that they were more concerned with the quality of the information, rather than where (or from whom) the information came. Finally, although this evidence is more indirect, in study 1 (chapter 4), participants preferred used a payoff bias which showed the best performing behaviour each round, rather than being influenced by the behaviour of a specific individual (prestige bias) or their group (conformity). These results may point towards individuals employing a content rather than a context bias (Laland, 2004), or, at minimum, that the content of the social information mattered more than the source. Previous research investigating content biases focused on different kinds of information than is considered here (Stubbersfield et al., 2017, 2018), meaning it is difficult to directly compare my results to previous research. However, previous experiments investigating the likeability of quotes (Acerbi & Tehrani, 2018) or story recall (Berl et al., 2021) have shown that the content of social information can be more influential than its source.

7.2.5. Summary of the main results

Taken together, these results suggest three things. Firstly, within a cooperative context, of the social learning strategies considered, it appears that a payoff bias is most influential, followed by conformity and then prestige bias. Secondly, given the choice, individuals appear to be selective in their sources of information to learn about cooperation, and were not highly influenced by the heuristic strategies (prestige bias and conformity) considered in this thesis. Finally, within cooperative contexts, the content of the social behaviour appears to be more relevant than the source of social behaviour.

7.3. Evaluating the impact of social learning

As described throughout section 7.2., this thesis did find some evidence that individuals used social learning or were influenced by social information within cooperative contexts. However, across all three empirical chapters, it was also the case that the effects associated with social learning were small in comparison to the other sources of influence on cooperative behaviour that I considered alongside social learning.

In study 1 (chapter 4), the agent-based simulation showed that the change in cooperation across rounds through social learning was minor compared to the variation in participant's cooperative tendencies (captured by the intercepts in the multilevel model). The cause of this variation in cooperative behaviour was not investigated in this study (see section 7.4.3.). Model comparison also only modestly favoured models which included the social learning strategies as predictors and their associated parameter values were small. A similar pattern was revealed in study 2 (chapter 5), where the social information presented to participants did not affect their decision to be altruistic or spiteful, as most participants opted to be altruistic. Finally, across study 3 (chapter 6) but particularly in the interviews, it was clear that participants held strong norms about the environment, which they felt were a stronger driver of their behaviour than the influence of others.

In the proceeding section, I will offer some potential explanations for the observed small effects of social learning on the whole (in addition to the factors discussed in the sections above). This will be driven by considering why an individual would use social learning and why this may not have been useful in the contexts considered in this thesis.

7.3.1. Asocial learning was not costly

Theory suggests that the principal adaptive purpose of social learning is the acquisition of traits while avoiding the costs of asocial learning (R. Kendal et al., 2018). Social learning is especially useful in scenarios where mistakes could be very costly, for example when choosing between novel food items. This trend is captured by state-based social learning strategies such as copy when uncertain or copy when the task is difficult (Laland, 2004), which have been demonstrated in experimental contexts in humans (Morgan et al., 2012; Toelch et al., 2010; Williamson & Meltzoff, 2011). When asocial learning is not costly, individuals rely less on social learning. I argue, it is questionable to what extent individuals were uncertain or perceived the task as difficult within the experimental cooperative contexts considered in this thesis, thus limiting the need for social learning.

Economic games were used in studies 1 and 2 (chapters 4 & 5) to assess cooperative (and spiteful) behaviour. It has been argued that economic games can induce confusion within participants, which in turn affects the interpretation of the results (Ferraro & Vossler, 2010). The confusion hypothesis was supported by a large-scale review, where the use of payoff learning was most consistent with participants starting confused and learning to improve their own payoffs (Burton-Chellew & West, 2021). Others reject the confusion hypothesis and instead suggest that humans are purposefully altruistic (Camerer, 2013). It follows that, if they are confused, participants are more likely to use social learning to work out what the optimal behaviour should be.

In the case of my experiments, I argue that participant confusion was not a significant factor. Primarily, this is because of the responses participants provided to my free text self-report questions. In both studies, out of a possible 10, the mean response was 8, which indicated that participants believed they understood how the game worked. The possible exception was study 2 (chapter 5), where some participants mistakenly believed they had behaved altruistically when they had been spiteful. Though, when combined with the high rating on the other self-report measure, it still appeared that participants perceived that they understood the game. Furthermore, the task required of participants (choosing a donation value) was not difficult and, although a participant's exact payoff in study 1 depended on the unknown behaviour of their groupmates, in both studies, participants could consistently earn the most on average by donating nothing (which was explained and shown to them visually prior to playing). Given this, there may have been no reason to use social learning within these studies, which may explain the relatively small observed effect of social information. Additionally, it was also not significantly costly to rely on asocial learning and participants may have considered (consciously or unconsciously) the costs of sub-optimal decision making to be too small to warrant an increased reliance on social learning.

The largest effects of social learning were found in relation to sustainability, where it is often difficult to know how one can act more sustainably and accurately assess the relative payoffs of alternative behaviours. In addition, asocial learning could also be costly, if individuals end up trying many expensive sustainable products. Consistent with this, participants in study 3 (chapter 6) and the literature more broadly have suggested knowledge as a barrier to their sustainability (Acharya, 2019; Axon, 2017; Young et al., 2010). Therefore, in a sustainability context, social learning may be more likely to be used than in experimental cooperative settings, because individuals are more likely to be uncertain.

7.3.2. The expected social norm was clear

A second use, or consequence, of social learning is the spread of social norms (Boyd & Richerson, 1985, 1994; Henrich & Boyd, 1998). As discussed in section 2.4.1., norms are often separated into descriptive (what do most people do) and injunctive (of what do most people approve) norms (Cialdini et al., 1991). This is relevant for cooperation, as it means that cooperative behaviour does not need to be copied directly; individuals could use social learning to acquire the expected norms surrounding cooperative behaviour. As such, participants could have used the behaviour of their groupmates to assess the normative behaviour within the experimental context. However, as discussed in section 7.2.2., there was little evidence that conformity (which is functionally similar to descriptive social norms) influenced participants' cooperative decision making in any of the studies. Injunctive norms were not explicitly considered in this thesis (discussed in section 7.4.1.), but the interview participants (study 3) generally believed that the influence of others was unconscious or in the form of broad social influences, rather than the direct copying of sustainable behaviour. Additionally, the strong personal norms they held towards sustainability were reported to be one of the strongest influences on their behaviour.

Using social learning to learn about cooperative social norms may not be necessary if the expected norm going into the experiment is very clear, or if individuals already possess a strong personal norm about cooperation in general. Across all three studies, there was an indication that this was the case. In study 1 (chapter 4), of the 136 participants that provided an interpretable response, 90 (66.18%) indicated that they believed the expected norm in the game to be cooperation. This aligned with the uncharacteristically high cooperation (compared to other public goods games (Zelmer, 2003)) that participants displayed in the experiment[19]. In study 2 (chapter 5), though the qualitative responses were less clear (40.11% indicated their behaviour was motivated by fairness or to benefit others), participants overall opted for altruism at much greater rates than neutral or spiteful behaviour. Although, it cannot be discounted that the self-report measures were themselves influenced by the social norms observed by participants in the experiment, this suggests that participants came into the experiment with some prior belief that cooperation was the expected social norm. Across both studies, participants' held cooperative norms may explain why social learning did not strongly influence participant's cooperative behaviour.

The social information observed in the experiment could also be thought of as an emergent social norm. Social norms are predicted to be most influential in contexts where there are opportunities for social evaluation, or with very clear expected norms (Gelfand & Harrington, 2015). In addition to the possibility that participants possessed an existing internal norm regarding cooperation, all the empirical research took place online (presenting little opportunity for social evaluation) which may have further limited the influence that social norms established during the experiment had on participants' behaviour. This is consistent with evidence that the impact of normative interventions is weakened in the face of strongly held intrinsically held social norms or other influential factors. In many sustainability interventions, normative interventions are less effective than those focusing on making sustainability more convenient (Carlson, 2001) or had little impact because the injunctive norm (recycling bottles) was already widely known and accepted (Bergquist et al., 2020). Cases where social norms are effective are those where it is plausible that the expected normative behaviour is not clear. For example, in a hotel where it is not obvious what hotel guests tend to do with used towels (Goldstein et al., 2008) or when queuing in a car where switching off the engine is not a widely accepted norm (Player et al., 2018). Other studies have shown that social norms (or other psychological factors) are less influential for sustainable behaviour when such behaviour is

[19] Though, this may have also been a consequence of the large multiplication factor used in this experiment (P. van den Berg et al., 2020)

convenient (Best & Kneip, 2011; Stern, 2011), indicating a further interaction with structural factors. This was supported by a model that showed while social learning can promote sustainable behaviour, individuals may switch behaviours if offered a more convenient alternative (Buenstorf & Cordes, 2008).

7.3.3. The consequence for cultural group selection

The apparently small effect of social learning within the context of cooperation in this thesis has implications for the role of cultural group selection to explain human cooperation. The basic tenant of cultural group selection is that, if multiple groups are competing for resources, the more cooperative groups will be more successful in these conflicts and, thus, cooperation will spread across the population (Boyd & Richerson, 2009a; Richerson et al., 2016). In this vein, cooperation is a cultural trait which can be maintained by mechanisms which stabilise group behaviour (D. Smith, 2020).

Recently, Smith, (2020) distinguished between two different modes of cultural group selection, normative and maladaptive. Under normative cultural group selection, cooperation is established as a group norm using costly punishment (Boyd & Richerson, 1992; Richerson et al., 2016). Under maladaptive cultural group selection, cooperative behaviour spreads through social learning (for example, through conformist or prestige biased transmission), without concern for the relative payoffs. The critical difference between these alternatives is that co-operators do not have reduced fitness under normative cultural group selection (as norm violators are punished) but they do when cooperation spreads through social learning.

Smith (2020) notes of the scant evidence for maladaptive cultural group selection, suggesting it is unlikely to account for the widespread cooperation observed in humans. The evidence presented in this thesis supports this general conclusion: individuals do not generally appear to adopt cooperative behaviour through social learning which, by extension, does not support maladaptive cultural group selection. I did, however find some evidence that social learning can spread uncooperative behaviour, which is consistent with previous experiments (Burton-Chellew, et al., 2017; Molleman et al., 2014).

This would suggest that the mechanism by which cultural group selection operates is more likely to be based on punishment and group norms, rather than individuals copying ultimately maladaptive behaviour.

7.4. Limitations and future directions

In this final section, I will outline some limitations of the work presented in this thesis. I will primarily, but not exclusively, describe methodological limitations including the kinds of social learning strategies investigated and the sampling methods employed. However, there are other potentially interesting theoretical limitations of the studies presented in this thesis, particularly with respect to factors relevant to cooperation and social learning that were not considered. I will detail four of the major limitations in turn and then discuss how they might inspire further research.

7.4.1. Other social learning strategies

As discussed in chapter 2 and throughout this thesis, I focused specifically on three of the major social learning strategies considered in the literature: payoff bias (McElreath & Henrich, 2003; Schlag, 1998), prestige bias (Henrich & Gil-White, 2001; Jiménez & Mesoudi, 2019) and conformity (Henrich & Boyd, 1998; Whiten, 2019a). To this end, all three empirical studies (besides, to some extent, the interviews) were designed and analysed with these strategies in mind. The statistical models used in study 1 (chapter 4) included only payoff, prestige and conformity parameters and experiment 2 (chapter 5) presented exclusively frequency or success weighted information to participants. Similarly, study 3 (chapter 6) restricted the vignettes to payoff, prestige and conformity framed information. Restricting my studies in this way was intentional and motivated by previous research (Burton-Chellew, el Mouden, et al., 2017; Henrich et al., 2015; Molleman et al., 2014; P. van den Berg, Molleman, & Weissing, 2015). My research was novel in that I explicitly considered the role of prestige and considered multiple social learning strategies simultaneously. However, this exposes the most obvious limitation of this thesis, the lack of a consideration of other social learning strategies.

There is mounting evidence that humans and animals use a variety of social learning strategies in different contexts (R. Kendal et al., 2018). Even within this thesis, some evidence for the use of other context (who) strategies than payoff, conformity and prestige was found using the semi-structured interviews. For instance, copying based on a model's proficiency (or knowledge) or familiarity. Wood et al., (2015) found some evidence that young children adopted a strategy demonstrated by a previously successful model, if their own innovated strategy matched that offered by model known to be unsuccessful in a similar task. School children who were given contradictory information about objects from familiar or unfamiliar teachers were more likely to endorse the familiar teacher's information, unless the teacher was seen to name other objects inaccurately (Corriveau & Harris, 2009).

It is also plausible that content biases (Laland, 2004) may be relevant for cooperative behaviour. Previous experiments have shown that socially relevant information is more likely to be transmitted and is better remembered by individuals (Mesoudi et al., 2006; Tehrani et al., 2015). Likewise, in a school classroom, gossip spread more quickly than sematic information (McGuigan & Cubillo, 2013). The findings from study 2 (chapter 5) indicated that the content of the social information was more important than the source. State based biases (discussed in section 7.3.1.) may also be relevant within a cooperative context. For example, in an experiment, participants induced with a good mood increased their cooperation more in response to a cooperative group norm than those induced with a bad mood (Hertel et al., 2000). More explicit consideration of content and state based social learning biases represents a ripe opportunity for further research.

The findings in chapter 6 pointed towards the relevance of familiar or knowledgeable models within the context of sustainability. It would therefore be interesting to employ interventions aiming to promote (real life) cooperation which vary information source and include cues of familiarity or knowledge, particularly contrasting domain based knowledge with a prestige cue such as status. Likewise, testing the role of familiar peers on lab-based cooperation may also prove fruitful. Perhaps by providing participants with information on a familiar individual's cooperation in an external group (as in other similar studies (Bardsley & Sausgruber, 2005; Burton-Chellew, el Mouden, et al., 2017;

Romano & Balliet, 2017)) could incentivise cooperation within the recipient of the information. The experimental studies in this thesis also did not explicitly test the role of injunctive norms. Given injunctive norms are often highly influential in applied contexts (Goldstein et al., 2008; Player et al., 2018; Schultz et al., 2007) and little evidence was found that participants directly copied cooperative behaviour, injunctive norms may also be influential in experimental cases.

Content biases are also well placed to help explain how cooperative reputations or knowledge of cooperative behaviour (relevant for indirect reciprocity (Nowak & Sigmund, 1998; Wu et al., 2016)) could spread through a population. A review of 300 years of newspaper articles showed that front page stories often feature altruism or cheating (Davis & McLeod, 2003) and contemporary newspapers continue to publish stories of this nature. To my knowledge, no study has systematically investigated whether cooperative or noncooperative behaviour is better remembered (but see Cosmides & Tooby, 1992) or whether a memory advantage translates into higher levels of cooperation. The reputational benefits of cooperative behaviour may also interact with model-based biases. In an experiment, participants were more cooperative towards richer than poorer individuals after seeing proportionally equivalent donations (Hackel & Zaki, 2018). Studies could investigate whether the identity of the model affects the kind of information that is most likely to be remembered and further explore the factors that affect the reputational benefit gained from cooperation.

7.4.2. The operationalisation of prestige and conformity

As described in the sections above (7.2.2./7.2.3.) the way the social learning strategies were operationalised in this thesis presents some limitations. The online experiments considered cooperation (and spite) on a continuum. This has many benefits, not least that it allows participants to express varying degrees of cooperative behaviour (Roberts & Sherratt, 1998) which exposes variability within participants' behaviour (Kurzban & Houser, 2001) but also that it offers greater ecological validity than arbitrarily restricting cooperation to a binary option. The downside is this complicates the consideration of conformity. To demonstrate evolutionary conformity, one must show that traits are copied disproportionately (Henrich & Boyd, 1998), however, it is unclear how this should be represented for continuous traits[20]. Consequently, most studies investigating conformity restrict behaviour to binary traits (Efferson et al., 2008; Haun et al., 2012; McElreath et al., 2008; Morgan et al., 2019). The studies presented in this thesis were essentially unable, by design, to demonstrate evolutionary conformity. They instead followed the approach of similar experiments and represented conformity using the mean (Bardsley & Sausgruber, 2005; Burton-Chellew, el Mouden, et al., 2017; Carpenter, 2004). Though little effect of conformity was found in this thesis regardless, this does introduce concerns of the extent that the results presented here can be compared to other studies investigating conformity and contextualised within cultural evolution more broadly.

The way that prestige was operationalised warrants further discussion. As discussed in section 7.2.3., though the cues of prestige employed in this thesis are consistent with theoretical definitions (Henrich, 2001; Jiménez & Mesoudi, 2019), they were not consistent between study 1 and study 3. The way prestige was indicated in study 1 also differed to some previous research (Atkisson et al., 2012; Chudek et al., 2012). This offers some benefits, as it allows a wider exploration of prestige and addresses the potential that some kinds of prestige cues (for example, popularity compared to general knowledge based) may have different influences on participant's behaviour. Nevertheless, considering this thesis did not support the strong theoretical evidence for the benefit of prestige on cooperation (Henrich et al., 2015; Lehmann & Feldman, 2008), this may be down to the specific ways in which prestige was defined in this thesis.

Despite the lack of strong effects for prestige and conformity, it may be useful for further research to address the limitations outlined above issues and continue to explore prestige and conformity. To investigate evolutionary conformity, it may be necessary to restrict cooperation to a binary decision, as it is considered in some models and experiments (Doebeli & Hauert, 2005a; Hilbig et al., 2018; Kümmerli et al., 2007) or consider a continuous trait within ordinal bins to allow an exploration of disproportionate copying. For prestige, studies could make use of cues such as looking times, which

[20] Morgan and Thompson (2020) represented conformity using the mean and variance of a trait in the population which agents updated their perception of based on their prior beliefs and the individuals they sampled from.

have successfully induced prestige effects in other contexts (Atkisson et al., 2012; Chudek et al., 2012). Alternatively, in two especially effective studies of this kind, Brand et al., (2020, 2021) used domain relevant quiz performance to signify prestige and found selective use of a prestige bias. Although this thesis found no effect associated with quiz performance, the methods employed by Brand et al., (2020, 2021) could be adapted to a cooperative context by investigating whether an individual's success in a similar but unrelated economic game affects the probability of them being copied. As my studies took place online, it is also worth considering whether conformity and prestige could be represented using cues more commonly present in an online setting. For example, the evidence for social learning biases was explored on Reddit by considering cues such as the content of the post, identity of the poster and the number of upvotes the post received (Priestley & Mesoudi, 2015). Given that the most compelling evidence for prestige comes from contexts where prestige cues emerged organically (7.2.3.) this may represent a more valid way to explore prestige in an online setting.

More generally, I suggest these points expose a need to assess the perception of social learning within participants. As described in section 7.2.3., the perceptions of prestige held by my interview participants were not always aligned with theoretical descriptions (Henrich & Gil-White, 2001). As such, although prestige may be theoretically important, it may not be practically relevant in real-life behaviour. Alternatively, it is also the case that individuals perceive social influence to have a smaller role on their behaviour than is true in reality (Frey & Meier, 2004b; Nolan et al., 2008). As such, expressed beliefs about the lack of influence of prestige does not invalidate its theoretical importance. Instead, this may point towards the possibility that these social learning strategies are not employed consciously (Heyes, 2016).

7.4.3. WEIRD, MTurk and student participants

Studies 1 and 2 (chapters 4 and 5) used the software Dallinger, (2020) to build and run the online experiments which automatically recruits participants from Amazon's Mechanical Turk (MTurk). The precise geographic region that can be recruited from MTurk varies with the time of day but is primarily US based. The use of MTurk offers many benefits, namely the speed and relatively low cost of recruiting participants and the diversity (relative to undergraduate students) of the sample (Aguinis et al., 2021). The drawbacks include participant drop out and inconsistency in English fluency (Aguinis et al., 2021). These issues affected studies 1 and 2 in this thesis. Study 1 (chapter 4) had several incomplete groups from participants dropping out (where data was discarded if the prestigious participants dropped out) and it was clear from some of the free text responses that some participants did not fully understand some aspects of the experiments. Although measures such as animated/interactive instructions and timed responses were put in place to manage these issues, this did not solve the issue entirely.

More broadly, many researchers have raised more general concerns about the use of MTurk, suggesting MTurk samples may be biased in some way (Aguinis et al., 2021). At the least, there is evidence that MTurk samples differ to traditional survey samples. On average, MTurk participants are younger, less likely to be female, higher educated and more politically liberal than samples drawn from highly representative social survey samples (Zack et al., 2019). Compared to student samples, MTurk participants were lower in extraversion, agreeableness and neuroticism (Colman et al., 2018) and showed larger variability on three political orientation scales to participants recruited using SONA (a participant recruitment system used by many Universities) or in person (Gamblin et al., 2017). However, while survey responses may differ between MTurk and other samples, experimental responses are more consistent between sampling methods (Coppock, 2019). Of relevance for this thesis, economic game results obtained on MTurk generally align with those from in person laboratory studies (Rand, 2012). Additionally, previous studies from cultural evolution have used Dallinger and recruited from MTurk and produced reliable results (Brand et al., 2020, 2021). One potential reason for the divide between the bias in survey and experimental responses is prior exposure of MTurk participants to similar (or identical) material (Aguinis et al., 2021).

Relatedly, concerns have been raised regarding the overreliance of many experimental studies on participants from WEIRD (western, educated, industrialised, rich, democratic) societies (Henrich, Heine, et al., 2010). The overwhelming majority of cooperation research (on which, most influential theories are based) has been conducted on student populations from WEIRD populations. A growing body of evidence is suggesting that cooperation and social learning use varies across cultures. For example, cooperation and "fair" behaviour varied across 15 societies according to their market integration (Henrich, Boyd, Bowles, et al., 2005; Henrich, Ensminger, et al., 2010) as did the use of antisocial punishment (Herrmann et al., 2008). Social learning also varies by cultural background (Mesoudi et al., 2014) and between societies (Lamba, 2014), as does the effect that country level sustainability norms have on individual's behaviour (Culiberg & Elgaaied-Gambier, 2016). Although MTurk may be a more diverse sample than students, it is nevertheless dominated by WEIRD participants and may be further biased towards those who have the technology and the necessary knowledge to access MTurk.

The scope of this thesis did not extend to cross cultural comparisons, which limits the conclusions that may be drawn from this thesis to some degree. The studies presented in chapter 6 were further limited in this respect, as they were conducted entirely with students. This presents two ways that the results and methods applied in this thesis can be taken forward. In addition to considering why populations may vary (discussed in section 7.4.4) further work is needed to better understand the extent of the variation in social learning and cooperative behaviour between cultural groups (Henrich, Heine, et al., 2010; Lamba, 2014; Mesoudi et al., 2014). Likewise, it would also be useful to expand the scope of the questionnaire and interviews to a different population (to offer a comparison with my student sample and offer analytical generalisability (B. Smith, 2018)) which in turn can inform the design of targeted interventions (Carattini et al., 2019; Rebele et al., 2021). For instance, the prediction that older individuals are less susceptible overall to social information (Flynn et al., 2016) could be evaluated by targeted interventions and similar methods to those presented in chapter 6.

7.4.4. Individual differences

Currently receiving theoretical attention is the variation present within and between cultural groups regarding cooperation (Henrich & Muthukrishna, 2021) and social learning (Mesoudi et al., 2016; Molleman et al., 2014). The use of social learning has been shown to vary alongside several other factors. The personality trait extraversion and overall centrality within social networks predicts the use of social learning (Rawlings et al., 2017) as does young age (Flynn et al., 2016). Of specific relevance to this thesis, variation in the use of prestige bias (Jiménez & Mesoudi, 2019) and conformity (Frey & Meier, 2004b; Muthukrishna et al., 2016) has also been shown. For example, older individuals were less affected by other's ratings when judging the attractiveness of faces (Little et al., 2015), indicating less use of prestige cues. Higher IQ was related to less social learning overall, but higher IQ individuals employed a stronger conformist bias than lower IQ individuals when social learning was used (Muthukrishna et al., 2016). Those primed with a positive mood increased their cooperation more in response to a cooperative group norm (Hertel et al., 2000). More generally, it has also been suggested that social learning preferences could, themselves, be a product of social or cultural influence (Heyes, 2018).

Participants in public goods games also vary in their degree of cooperation; most participants can be grouped into strong co-operators, free riders or reciprocators (Kurzban & Houser, 2001) and they additionally vary in the way that they respond to the contributions of others (P. van den Berg, Molleman, Junikka, et al., 2015). Like social learning, cooperation additionally varies based on other factors such as high agreeableness (Hirsh, 2014) and emotions such as outrage (van de Vyver & Abrams, 2015). There is additional variation in the type of social information participants use with cooperative games (Lamba, 2014; P. van den Berg, Molleman, & Weissing, 2015) where those that use conformity are more cooperative than payoff-based learners (P. van den Berg, Molleman, & Weissing, 2015).

This thesis was limited in that it did not explicitly consider the role of individual differences, despite finding indirect evidence of it across all three studies. Study 1 (chapter 4) found considerable variation in participant's cooperative behaviour, which was more influential than social learning. Likewise, the interview responses indicated variation in the extent that individuals used or were influenced by social information. What was not possible (beyond speculation) from the data collected in this thesis, was to address how and why individuals varied. Furthermore, the design of study 2 (chapter 5) and the analysis of study 1 (chapter 4) did not account for the possibility that participants may differ in their sensitivity to the social information or vary in their use of social learning (R. Kendal et al., 2018).

Theoretically, understanding the variation in how cooperative behaviour may spread through a group is a vital aspect of cultural group selection (Richerson et al., 2016) or for the transmission of cultural traits more generally. Practically, understanding when and why individuals are likely to be influenced by particular kinds of (social) information is also important for the design of behaviour-change interventions. Individuals who are altruistically motivated may respond to interventions which highlight the benefits of their cooperation or the expected cooperation of others, while those that are self-interested would be more likely to respond to interventions emphasising the personal benefits of their behaviour (Burton-Chellew & West, 2021). Relatedly, others have suggested that interventions should target those most resistant to change because those less strongly inclined towards the behaviour will follow the shifted social norm (Efferson et al., 2020). This is in addition to the differential use and effects of social learning across cultures (section 7.4.3.) which may further influence the kinds of interventions that may be effective.

The approaches taken in this thesis could be expanded to address these limitations. Firstly, greater attention should be given to the extent, and source, of variation between individuals in the use of social learning or cooperative behaviour. This may be based on personality (Hirsh, 2014; Rawlings et al., 2017), their motivation for cooperative behaviour (Burton-Chellew & West, 2021), tendency to use social learning (R. Kendal et al., 2018) or more generalised factors that moderate the extent that individuals are influenced by social norms (Gelfand & Harrington, 2015). Such findings could then be used in an applied context to help promote real life cooperative behaviour. Rebele et al., (2021) discuss the relevance of factors such as personality for the design of interventions, which can help inform what behaviours individuals are disposed towards and, thus, what kinds of interventions are likely to be effective. It may also be the case that community structure may influence the kinds of social learning and patterns of social learning exhibited. For example, Lamba, (2014) found conformity to be more common in larger communities where individuals had more social connections.

The analysis methods in study 1 (chapter 4) and the design of studies 1 and 2 could be meaningfully expanded to address the variation in participants use of social learning. Across both experiments, allowing participants to choose the social information they view (as in (Flynn et al., 2016; P. van den Berg, Molleman, & Weissing, 2015)) rather than presenting it based on the experimental condition would permit an investigation of who prefers different kinds of social information and why. Additionally, a more accurate measure of the influence of a given strategy is possible, as effect sizes would not be confounded by the inclusion of individuals that are uninterested in particular types of social information. Relatedly, repeated observations of the same individuals mean multilevel modelling can estimate a social learning parameter for each individual (Toyokawa et al., 2019) (rather than a single pooled estimate as in chapter 4). Then, participants can be grouped according to the social learning strategy they tended to use, based on the strength of the slope parameter (Toyokawa et al., 2019). Practical and financial limitations on this thesis meant it was not possible to collect a sufficiently large sample to apply this method.

7.4.5. Punishment and cooperative payoff structure

As discussed in chapter 2 (2.1.2.4.) costly punishment is argued to be an important mechanism by which cooperative behaviour can be maintained within a population (Boyd & Richerson, 1992; Boyd et al., 2003) and additionally important for cultural group selection (Richerson et al., 2016; D. Smith, 2020). This is especially relevant for cultural group selection, as the punishment of non-co-operators is one of the proposed mechanisms by which cooperation can be maintained between social groups. Furthermore, there is also evidence that punishment may be socially transmitted. Two experiments have shown that individuals are encouraged to use punishment based on the expressed preferences of third-party individuals (FeldmanHall et al., 2018) and in response to group norms promoting punishment or (to a lesser extent) cooperation (Li et al., 2021). Few experimental studies have emulated the conditions necessary for cultural group selection, for example by including competition between groups (Majolo & Maréchal, 2017; Puurtinen & Mappes, 2009b) alongside punishment (Balliet, Mulder, et al., 2011). There is also no reason apriori to assume that punishment could not also occur alongside the social learning of cooperative behaviour. Study 1 (chapter 4) could be expanded to include punishment and competition between groups. Though such a study would be ambitious, the relative ease and inexpensive nature of online experiments may allow for this possibility.

Relatedly, it may also be interesting to investigate other game structures than the ones considered in this thesis (prisoner's dilemma and snowdrift). For example, previous studies have considered the coordination or stag hunt game (Molleman et al., 2014; P. van den Berg, Molleman, & Weissing, 2015) and found that the kind of social learning used by participant varies between them. This may be achieved by changing the underlying payoff structure of the experimental game. Alternatively, as some researchers have cautioned against assuming that economic games apply to real life behaviour (Levitt & List, 2007), it may be useful to consider a cooperative scenario beyond a typical economic game (for example, Hertel et al., 2000; Madsen et al., 2007) or make greater use of field-based studies. Along these lines, an initial experimental condition for study 3 which was not used in this thesis was a comparison of intervention materials that framed the cooperative behaviour as a PD (emphasising the altruistic nature of sustainability) or a SD game (emphasising the personal benefits irrespective of the behaviour of others). As described in section 7.4.4., such framing may appeal to different types of people (Burton-Chellew & West, 2021).

7.5. Concluding remarks

To sum up, through the studies presented in this thesis, I found some evidence that participants used social learning or were influenced by social information in cooperative scenarios. Across the three empirical chapters, a clear hierarchy of influence emerged between the three social learning strategies considered. Payoff weighted information (payoff bias) was the most influential, then frequency information (conformity) and finally prestige weighted information (prestige). However, the effects of social learning were small and less influential than other factors. More generally, the studies in this thesis employed novel methods and offered broader insights into the dynamics of cooperation between prisoner's dilemma and snowdrift games and explored costly spiteful behaviour. Further research could beneficially continue to explore the role of social learning on cooperative behaviour, giving particular attention to individual and cultural differences and continue to find ways that these insights can be applied to real life cooperative dilemmas.

References

Aarts, H., & Dijksterhuis, A. (2003). The Silence of the Library: Environment, Situational Norm, and Social Behavior. *Journal of Personality and Social Psychology*, *84*(1), 18–28. https://doi.org/10.1037/0022-3514.84.1.18

Abbink, K., & Herrmann, B. (2011). The moral costs of nastiness. *Economic Inquiry*, *49*(2), 631–633. https://doi.org/10.1111/j.1465-7295.2010.00309.x

Abbink, K., & Sadrieh, A. (2009). The pleasure of being nasty. *Economics Letters*, *105*(3), 306–308. https://doi.org/10.1016/j.econlet.2009.08.024

Abrahamse, W., & Steg, L. (2013). Social influence approaches to encourage resource conservation: A meta-analysis. *Global Environmental Change*, *23*(6), 1773–1785. https://doi.org/10.1016/j.gloenvcha.2013.07.029

Acerbi, A., Mesoudi, A., & Smollo Marco. (2020). *Individual-based models of cultural evolution. A step-by-step guide using R*. https://doi.org/doi:110.31219/osf.io/32v6a

Acerbi, A., & Tehrani, J. J. (2018). Did Einstein Really Say that? Testing Content Versus Context in the Cultural Selection of Quotations. *Journal of Cognition and Culture*, *18*(3–4), 293–311. https://doi.org/10.1163/15685373-12340032

Acerbi, A., van Leeuwen, E. J. C., Haun, D. B. M., & Tennie, C. (2016). Conformity cannot be identified based on population-level signatures. *Scientific Reports*, *6*(1), 1–9. https://doi.org/10.1038/srep36068

Acharya, A. (2019). Factors behind dissuasion to green products among young consumers: A qualitative study. *Qualitative Report*, *24*(12), 3197–3214.

Achtnicht, M. (2011). Do environmental benefits matter? Evidence from a choice experiment among house owners in Germany. *Ecological Economics*, *70*(11), 2191–2200. https://doi.org/10.1016/j.ecolecon.2011.06.026

Aguinis, H., Villamor, I., & Ramani, R. S. (2021). MTurk Research: Review and Recommendations. *Journal of Management*, *47*(4), 823–837. https://doi.org/10.1177/0149206320969787

Akçay, Çağlar, Reed, V. A., Campbell, S. E., Templeton, C. N., & Beecher, M. D. (2010). Indirect reciprocity: Song sparrows distrust aggressive neighbours based on eavesdropping. *Animal Behaviour*, *80*(6), 1041–1047. https://doi.org/10.1016/j.anbehav.2010.09.009

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, *95*(9–10), 1082–1095. https://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. https://doi.org/10.1016/j.econlet.2009.05.016

Anderson, L. R., Mellor, J. M., & Milyo, J. (2008). Inequality and public good provision: An experimental analysis. *Journal of Socio-Economics*, *37*(3), 1010–1028. https://doi.org/10.1016/j.socec.2006.12.073

Andersson, E., & Öhman, J. (2017). Young people’s conversations about environmental and sustainability issues in social media. *Environmental Education Research*, *23*(4), 465–485. https://doi.org/10.1080/13504622.2016.1149551

Andor, M. A., & Fels, K. M. (2018). Behavioral Economics and Energy Conservation – A Systematic Review of Non-price Interventions and Their Causal Effects. *Ecological Economics*, *148*(January), 178–210. https://doi.org/10.1016/j.ecolecon.2018.01.018

Andras, P. (2018). Social learning in repeated cooperation games in uncertain environments. *Cognitive Systems Research*, *51*, 24–39. https://doi.org/10.1016/j.cogsys.2018.04.013

Andreoni, J. (1988). Why free ride?. Strategies and learning in public goods experiments. *Journal of Public Economics*, *37*, 291–304. https://doi.org/10.1016/0047-2727(88)90043-6

Andreoni, J. (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal*, *100*, 464–477. https://doi.org/10.2307/2234133

Andrés Guzmán, R., Rodríguez-Sickert, C., & Rowthorn, R. (2007). When in Rome, do as the Romans do: the coevolution of altruistic punishment, conformist learning, and cooperation. *Evolution and Human Behavior*, *28*(2), 112–117. https://doi.org/10.1016/j.evolhumbehav.2006.08.002

Apicella, C. L., Marlowe, F. W., Fowler, J. H., Nicholas, A., Diego, S., & Diego, S. (2012). Social Networks and Cooperation in Hunter-Gatherers. *Nature*, *481*(7382), 497–501. https://doi.org/10.1038/nature10736.Social

Aplin, L. M. (2019). Culture and cultural evolution in birds: a review of the evidence. *Animal Behaviour*, *147*, 179–187. https://doi.org/10.1016/j.anbehav.2018.05.001

Armel, K. C., Yan, K., Todd, A., & Robinson, T. N. (2011). The Stanford Climate Change Behavior Survey (SCCBS): Assessing greenhouse gas emissions-related behaviors in individuals and populations. *Climatic Change*, *109*(3–4), 671–694. https://doi.org/10.1007/s10584-011-0031-y

Arseneau-Robar, T. J. M., Taucher, A. L., Müller, E., van Schaik, C., Bshary, R., & Willems, E. P. (2016). Female monkeys use both the carrot and the stick to promote male participation in intergroup fights. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1843), 20161817. https://doi.org/10.1098/rspb.2016.1817

Asensio, O. I., & Delmas, M. A. (2015). Nonprice incentives and energy conservation. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(6), E510–E515. https://doi.org/10.1073/pnas.1401880112

Atkisson, C., O'Brien, M. J., & Mesoudi, A. (2012). Adult learners in a novel environment use prestige-biased social learning. *Evolutionary Psychology, 10*(3), 519–537. https://doi.org/10.1177/147470491201000309

Axelrod, R., & Hamilton, W. D. (1981). The Evolution of Cooperation. *Evolution, 211*(4489), 1390–1396. https://doi.org/10.1086/383541

Axon, S. (2017). "Keeping the ball rolling": Addressing the enablers of, and barriers to, sustainable lifestyles. *Journal of Environmental Psychology, 52*, 11–25. https://doi.org/10.1016/j.jenvp.2017.05.002

Baillon, A., Selim, A., & van Dolder, D. (2013). On the social nature of eyes: The effect of social cues in interaction and individual choice tasks. *Evolution and Human Behavior, 34*(2), 146–154. https://doi.org/10.1016/j.evolhumbehav.2012.12.001

Baldini, R. (2013). Two success-biased social learning strategies. *Theoretical Population Biology, 86*, 43–49. https://doi.org/10.1016/j.tpb.2013.03.005

Balliet, D., Li, N. P., Macfarlan, S. J., & Van Vugt, M. (2011). Sex Differences in Cooperation: A Meta-Analytic Review of Social Dilemmas. *Psychological Bulletin, 137*(6), 881–909. https://doi.org/10.1037/a0025354

Balliet, D., Mulder, L. B., & van Lange, P. A. M. (2011). Reward, punishment, and cooperation: A meta-analysis. *Psychological Bulletin, 137*(4), 594–615. https://doi.org/10.1037/a0023489

Bandura, A. (1986). *Social Foundations of Thought and Action*. Prentice-Hall.

Bangerter, A., & Heath, C. (2004). The Mozart effect: Tracking the evolution of a scientific legend. *British Journal of Social Psychology, 43*(4), 605–623. https://doi.org/10.1348/0144666042565353

Banks, S. C., Lindenmayer, D. B., Mcburney, L., Blair, D., Knight, E. J., & Blyton, M. D. J. (2011). Kin selection in den sharing develops under limited availability of tree hollows for a forest marsupial. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1719), 2768–2776. https://doi.org/10.1098/rspb.2010.2657

Barclay, P., & Krupp, D. B. (2016). The burden of proof for a cultural group selection account. In *Behavioral and Brain Sciences*. https://doi.org/10.1017/S0140525X15000060

Bardsley, N., & Sausgruber, R. (2005). Conformity and reciprocity in public good provision. *Journal of Economic Psychology*, *26*(5), 664–681. https://doi.org/10.1016/j.joep.2005.02.001

Barrett, B. J., McElreath, R. L., & Perry, S. E. (2017). Pay-off-biased social learning underlies the diffusion of novel extractive foraging traditions in a wild primate. *Proceedings of the Royal Society B: Biological Sciences*, *284*, 20170358. https://doi.org/10.1098/rspb.2017.0358

Barron, A. B., Oldroyd, B. P., & Ratnieks, F. L. W. (2001). Worker reproduction in honey-bees (Apis) and the anarchic syndrome: A review. *Behavioral Ecology and Sociobiology*, *50*, 199–208. https://doi.org/10.1007/s002650100362

Bartlett, F. C. (1932). Remembering: A study in experimental and social psychology. In *Learning and Behavior*. Cambridge University Press. https://doi.org/10.3758/bf03196002

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, *2*(3), 412–414. https://doi.org/10.1098/rsbl.2006.0509

BBC News. (2011). *London bombings: unsung heroes of 7 July*. file:///Users/Gary/Desktop/Gary’s Library/Library.papers3/Articles/1999/Unknown/1999-1%5Cnpapers3://publication/uuid/4D88F43C-896A-4C69-B310-23AA660007B9

BBC News. (2021). *Capt Tom donations: What was the £33m spent on?*

Bebbington, K., MacLeod, C., Ellison, T. M., & Fay, N. (2017). The sky is falling: evidence of a negativity bias in the social transmission of information. *Evolution and Human Behavior*, *38*(1), 92–101. https://doi.org/10.1016/j.evolhumbehav.2016.07.004

Bedard, S. A. N., & Tolmie, C. R. (2018). Millennials' green consumption behaviour: Exploring the role of social media. *Corporate Social Responsibility and Environmental Management*, *25*(6), 1388–1396. https://doi.org/10.1002/csr.1654

Beheim, B. A., Thigpen, C., & Mcelreath, R. (2014). Strategic social learning and the population dynamics of human behavior: The game of Go. *Evolution and Human Behavior*, *35*(5), 351–357. https://doi.org/10.1016/j.evolhumbehav.2014.04.001

Benz, M., & Meier, S. (2008). Do people behave in experiments as in the field?-Evidence from donations. *Experimental Economics*, *11*(3), 268–281. https://doi.org/10.1007/s10683-007-9192-y

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, *10*, 122–142. https://doi.org/10.1006/game.1995.1027

Berger, J., & Milkman, K. (2010). Social Transmission, Emotion, and the Virality of Online Content. *Wharton Research Paper*, *10*, 1–53.

Bergquist, M., & Nilsson, A. (2016). I saw the sign: Promoting energy conservation via normative prompts. *Journal of Environmental Psychology*, *46*, 23–31. https://doi.org/10.1016/j.jenvp.2016.03.005

Bergquist, M., Nyström, L., & Nilsson, A. (2020). Feeling or following? A field-experiment comparing social norms-based and emotions-based motives encouraging pro-environmental donations. *Journal of Consumer Behaviour*, *19*(4), 351–358. https://doi.org/10.1002/cb.1813

Berl, R. E. W., Samarasinghe, A. N., Roberts, S. G., Jordan, F. M., & Gavin, M. C. (2021). Prestige and content biases together shape the cultural transmission of narratives. *Evolutionary Human Sciences*, *3*(e42), 1–22. https://doi.org/10.1017/ehs.2021.37

Bernasconi, G., & Strassmann, J. E. (1999). Cooperation among unrelated individuals: The ant foundress case. *Trends in Ecology and Evolution*, *14*(12), 477–482. https://doi.org/10.1016/S0169-5347(99)01722-X

Best, H., & Kneip, T. (2011). The impact of attitudes and behavioral costs on environmental behavior: A natural experiment on household waste recycling. *Social Science Research*, *40*(3), 917–930. https://doi.org/10.1016/j.ssresearch.2010.12.001

Bhogal, M. S., Farrelly, D., & Galbraith, N. (2019). The role of prosocial behaviors in mate choice: A critical review of the literature. *Current Psychology*, *38*(4), 1062–1075. https://doi.org/10.1007/s12144-019-00308-8

Bicchieri, C. (2005). The grammar of society: The nature and dynamics of social norms. In *The Grammar of Society: The Nature and Dynamics of Social Norms*. https://doi.org/10.1017/CBO9780511616037

Binmore, K. (2007a). *Playing for Real: A Text on Game Theory*. Oxford University Press. https://doi.org/10.1007/s00712-007-0304-0

Binmore, K. (2007b). Playing for Real: A Text on Game Theory. In *Mental health today*. https://doi.org/10.1007/s00712-007-0304-0

Blackwell, C., & Diamond, Z. (2017). Combatting the Joy of Destruction with Pro-Social Behavior. *Review of Behavioural Economics*, *4*(3), 275–293. https://doi.org/10.2139/ssrn.2795625

Blok, V., Wesselink, R., Studynka, O., & Kemp, R. (2015). Encouraging sustainability in the workplace: A survey on the pro-environmental behaviour of university employees. *Journal of Cleaner Production*, *106*, 55–67. https://doi.org/10.1016/j.jclepro.2014.07.063

Boesch, C., & Boesch, H. (1989). Hunting behavior of wild chimpanzees in the Taï National Park. *American Journal of Physical Anthropology*, *78*(4), 547–573. https://doi.org/10.1002/ajpa.1330780410

Bolle, F., & Spiller, J. (2021). Cooperation against all predictions. *Economic Inquiry*, *59*(3), 904–924. https://doi.org/10.1111/ecin.12976

Bond, R., & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using asch's (1952b, 1956) line judgment task. *Psychological Bulletin*, *119*(1), 111–137. https://doi.org/10.1037/0033-2909.119.1.111

Bono, A. E. J., Whiten, A., van Schaik, C., Krützen, M., Eichenberger, F., Schnider, A., & van de Waal, E. (2018). Payoff- and Sex-Biased Social Learning Interact in a Wild Primate Population. *Current Biology*, *28*(17), 2800-2805.e4. https://doi.org/10.1016/j.cub.2018.06.015

Bowles, S. (2008). Policies designed for self-interested citizens may undermine "The moral sentiments": Evidence from economic experiments. *Science*, *320*(5883), 1605–1609. https://doi.org/10.1126/science.1152110

Boyd, R., Gintis, H., & Bowles, S. (2010). Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science*, *328*(5978), 617–620. https://doi.org/10.1126/science.1183665

Boyd, R., & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. University of Chicago Press. https://doi.org/10.1097/00005053-198702000-00018

Boyd, R., & Richerson, P. J. (1988). The evolution of reciprocity in sizable groups. *Journal of Theoretical Biology*, *132*, 337–356. https://doi.org/10.1016/S0022-5193(88)80219-4

Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, *13*(3), 171–195. https://doi.org/10.1016/0162-3095(92)90032-Y

Boyd, R., & Richerson, P. J. (1994). The evolution of norms: an anthropological view. *Journal of Institutional and Theoretical Economics*, *150*(1), 72–87. https://doi.org/10.2307/40753018

Boyd, R., & Richerson, P. J. (1995). Why does culture increase human adaptability? *Ethology and Sociobiology*, *16*, 125–143. https://doi.org/10.1016/0162-3095(94)00073-G

Boyd, R., & Richerson, P. J. (2005). *The Origin and Evolution of Cultures*. Oxford University Press. https://doi.org/10.1186/jbiol139

Boyd, R., & Richerson, P. J. (2009a). Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1533), 3281–3288. https://doi.org/10.1098/rstb.2009.0134

Boyd, R., & Richerson, P. J. (2009b). Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1533), 3281–3288. https://doi.org/10.1098/rstb.2009.0134

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(SUPPL. 2), 10918–10925. https://doi.org/10.1073/pnas.1100290108

Boyd, Robert., Gintis, H., Bowles, S., & Richerson, P. J. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, *100*(6), 3531–3535. https://doi.org/10.1073/pnas.0630443100

Braeken, J., & van Assen, M. A. L. M. (2017). An empirical Kaiser criterion. *Psychological Methods*, *22*(3), 450–466. https://doi.org/10.1037/met0000074

Brand, C. O., Heap, S., Morgan, T. J. H., & Mesoudi, A. (2020). The emergence and adaptive use of prestige in an online social learning task. *Scientific Reports*, *10*(1), 1–11. https://doi.org/10.1038/s41598-020-68982-4

Brand, C. O., & Mesoudi, A. (2019). Prestige and dominance-based hierarchies exist in naturally occurring human groups, but are unrelated to task-specific knowledge. *Royal Society Open Science*, *6*(5). https://doi.org/10.1098/rsos.181621

Brand, C. O., Mesoudi, A., & Morgan, T. J. H. (2021). Trusting the experts: The domain-specificity of prestige-biased social learning. *PLoS ONE*, *16*(8 August), 1–15. https://doi.org/10.1371/journal.pone.0255346

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2013). *Successful Qualitative Research: A Practical Guide for Beginners*. Sage Publications.

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Brereton, A. R. (1994). Return-benefit spite hypothesis: An explanation for sexual interference in stumptail macaques (Macaca arctoides). *Primates*, *35*(2), 123–136. https://doi.org/10.1007/BF02382049

Brooks, J. S., Waring, T., Borgerhoff Mulder, M., & Richerson, P. J. (2018). Applying cultural evolution to sustainability challenges: an introduction to the special issue. *Sustainability Science*, *13*(1), 1–8. https://doi.org/10.1007/s11625-017-0516-3

Brown, J. S., & Vincent, T. L. (2008). Evolution of cooperation with shared costs and benefits. *Proceedings of the Royal Society B: Biological Sciences*, *275*, 1985–1994. https://doi.org/10.1098/rspb.2007.1685

Bruhin, A., Janizzi, K., & Thöni, C. (2020). Uncovering the heterogeneity behind cross-cultural variation in antisocial punishment. *Journal of Economic Behavior and Organization*, *180*, 291–308. https://doi.org/10.1016/j.jebo.2020.10.005

Bryan, J. H., & Test, M. A. (1967). Models and Helping: Naturalistic Studies in Aiding Behavior. *Journal of Personality and Social Psychology*, *6*(4 PART 1), 400–407. https://doi.org/10.1037/h0024826

Bryman, A. (2004). Social Research Methods. In *Social Research methods* (Vol. 2nd). Oxford University Press. https://doi.org/10.4135/9781849209939

Bshary, R., & Grutter, A. S. (2002). Asymmetric cheating opportunities and partner control in a cleaner fish mutualism. *Animal Behaviour*, *63*(3), 547–555. https://doi.org/10.1006/anbe.2001.1937

Bshary, R., & Grutter, A. S. (2006). Image scoring and cooperation in a cleaner fish mutualism. *Nature*, *441*(7096), 975–978. https://doi.org/10.1038/nature04755

Buckley, E., & Croson, R. (2006). Income and wealth heterogeneity in the voluntary provision of linear public goods. *Journal of Public Economics*, *90*(4–5), 935–955. https://doi.org/10.1016/j.jpubeco.2005.06.002

Buenstorf, G., & Cordes, C. (2008). Can sustainable consumption be learned? A model of cultural evolution. *Ecological Economics*, *67*(4), 646–657. https://doi.org/10.1016/j.ecolecon.2008.01.028

Bulbulia, J. (2004). Religious Costs as Adaptations. *Evolution and Cognition*, *10*(1), 19–42.

Bulbulia, J., & Sosis, R. (2011). Signalling theory and the evolution of religious cooperation. *Religion*, *41*(3), 363–388. https://doi.org/10.1080/0048721X.2011.604508

Burger, J. M., Messian, N., Patel, S., del Prado, A., & Anderson, C. (2004). What a Coincidence! The Effects of Incidental Similarity on Compliance. *Personality and Social Psychology Bulletin*, *30*(1), 35–43. https://doi.org/10.1177/0146167203258838

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Bürkner, P. C. (2021). *Handle Missing Values with brms*. https://cran.r-project.org/web/packages/brms/vignettes/brms_missings.html#compatibility-with-other-multiple-imputation-packages

Burn, S. M. (1991). Social Psychology and the Stimulation of Recycling Behaviors: The Block Leader Approach. *Journal of Applied Social Psychology*, *21*(8), 611–629. https://doi.org/10.1111/j.1559-1816.1991.tb00539.x

Burton-Chellew, M. N., & Amico, V. D. (2021). A preference to learn from successful rather than common behaviours in human social dilemmas. *Proceedings of the Royal Society B: Biological Sciences, 288*, 20211590.

Burton-Chellew, M. N., El Mouden, C., & West, S. A. (2017). Evidence for strategic cooperation in humans. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1856). https://doi.org/10.1098/rspb.2017.0689

Burton-Chellew, M. N., el Mouden, C., & West, S. A. (2017). Social learning and the demise of costly cooperation in humans. *Proceedings of the Royal Society of London B: Biological Sciences*, *284*(1853), 10–12. https://doi.org/10.1098/rspb.2017.0067

Burton-Chellew, M. N., Nax, H. H., & West, S. A. (2015). Payoff-based learning explains the decline in cooperation in public goods games. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1801), 20142678–20142678. https://doi.org/10.1098/rspb.2014.2678

Burton-Chellew, M. N., & West, S. A. (2013). Prosocial preferences do not explain human cooperation in public-goods games. *Proceedings of the National Academy of Sciences*, *110*(1), 216–221. https://doi.org/10.1073/pnas.1210960110

Burton-Chellew, M. N., & West, S. A. (2021). Payoff-based learning best explains the rate of decline in cooperation across 237 public-goods games. *Nature Human Behaviour*, *5*, 36–38. https://doi.org/10.1038/s41562-021-01107-7

Büyükboyaci, M. (2014). Risk attitudes and the stag-hunt game. *Economics Letters*, *124*(3), 323–325. https://doi.org/10.1016/j.econlet.2014.06.019

Camerer, C. F. (2013). Experimental, cultural, and neural evidence of deliberate prosociality. *Trends in Cognitive Sciences*, *17*(3), 106–108. https://doi.org/10.1016/j.tics.2013.01.009

Capraro, V., Rodriguez-Lara, I., & Ruiz-Martos, M. J. (2020). Preferences for efficiency, rather than preferences for morality, drive cooperation in the one-shot Stag-Hunt game. *Journal of Behavioral and Experimental Economics* , *86*(January), 101535. https://doi.org/10.1016/j.socec.2020.101535

Carattini, S., Levin, S., & Tavoni, A. (2019). Cooperation in the Climate Commons. *Review of Environmental Economics and Policy*, *13*(2), 227–247. https://doi.org/10.1093/reep/rez009

Carlson, A. E. (2001). Recycling Norms. *California Law Review*, *89*(5), 1231–1300.

Carpenter, J. P. (2004). When in Rome: Conformity and the provision of public goods. *Journal of Socio-Economics*, *33*(4), 395–408. https://doi.org/10.1016/j.socec.2004.04.009

Carpenter, J. P., Matthews, P. H., & Ong'ong'a, O. (2004). Why punish? Social reciprocity and the enforcement of prosocial norms. *Journal of Evolutionary Economics*, *14*(4), 407–429. https://doi.org/10.1007/s00191-004-0212-1

Carter, G. G., & Wilkinson, G. S. (2013). Food sharing in vampire bats: Reciprocal help predicts donations more than relatedness or harassment. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1753). https://doi.org/10.1098/rspb.2012.2573

Cartwright, E., Gillet, J., & Van Vugt, M. (2013). Leadership by example in the weak-link game. *Economic Inquiry*, *51*(4), 2028–2043. https://doi.org/10.1111/ecin.12003

Casey, P. J., & Scott, K. (2006). Environmental concern and behaviour in an Australian sample within an ecocentric - Anthropocentric framework. *Australian Journal of Psychology*, *58*(2), 57–67. https://doi.org/10.1080/00049530600730419

Chai, B. C., van der Voort, J. R., Grofelnik, K., Eliasdottir, H. G., Klöss, I., & Perez-Cueto, F. J. A. (2019). Which diet has the least environmental impact on our planet? A systematic review of vegan, vegetarian and omnivorous diets. *Sustainability (Switzerland)*, *11*(15). https://doi.org/10.3390/su11154110

Chen, Y. S., & Chang, C. H. (2012). Enhance green purchase intentions: The roles of green perceived value, green perceived risk, and green trust. *Management Decision*, *50*(3), 502–520. https://doi.org/10.1108/00251741211216250

Christman, B. (2013). A Brief History Of Environmental Law In The UK. In *Environmental Scientist* (Vol. 22, Issue 4).

Chuah, S. H., Hoffmann, R., Ramasamy, B., & Tan, J. H. W. (2014). Religion, ethnicity and cooperation: An experimental study. *Journal of Economic Psychology*, *45*, 33–43. https://doi.org/10.1016/j.joep.2014.07.002

Chudek, M., Baron, A. S., & Birch, S. (2016). Unselective Overimitators: The Evolutionary Implications of Children's Indiscriminate Copying of Successful and Prestigious Models. *Child Development*, *87*(3), 782–794. https://doi.org/10.1111/cdev.12529

Chudek, M., Heller, S., Birch, S., & Henrich, J. (2012). Prestige-biased cultural learning: bystander's differential attention to potential models influences children's learning. *Evolution and Human Behavior*, *33*(1), 46–56. https://doi.org/10.1016/j.evolhumbehav.2011.05.005

Chwialkowska, A. (2019). How sustainability influencers drive green lifestyle adoption on social media: The process of green lifestyle adoption explained through the ;enses of the minority influence model and social learning theory. *Management of Sustainable Development*, *11*(1), 33–43.

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. https://doi.org/10.1016/S0065-2601(08)60330-5

Cialdini, R. B., Trost, M. R., Reno, R. R., & Kallgren, C. a. (1998). Social influence: Social norms, conformity and compliance. In *The handbook of social psychology, Vols. 1 and 2*. https://doi.org/10.1037/0022-3514.58.6.1015

Claidière, N., Bowler, M., Brookes, S., Brown, R., & Whiten, A. (2014). Frequency of behavior witnessed and conformity in an everyday social context. *PLoS ONE*, *9*(6), 1–10. https://doi.org/10.1371/journal.pone.0099874

Claidiere, N., & Whiten, A. (2012). Integrating the study of conformity and culture in humans and nonhuman animals. *Psychological Bulletin*, *138*(1), 126–145. https://doi.org/10.1037/a0025868

Clutton-Brock, T. H. (2009). Cooperation between non-kin in animal societies. *Nature*, *462*(7269), 51–57. https://doi.org/10.1038/nature08366

Clutton-Brock, T. H., Russell, A. F., Sharpe, L. L., Brotherton, P. N. M., McIlrath, G. M., White, S., & Cameron, E. Z. (2001). Effects of helpers on juvenile development and survival in meerkats. *Science*, *293*(5539), 2446–2449. https://doi.org/10.1126/science.1061274

Cohen, E. (2012). The Evolution of Tag-Based Cooperation in Humans. *Current Anthropology*, *53*(5), 588–616. https://doi.org/10.1086/667654

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Cohen, J. (1995). The earth is round (p < .05). *American Psychologist*, *50*(12), 1103–1103. https://doi.org/10.1037/0003-066X.50.12.1103

Colman, D. E., Vineyard, J., & Letzring, T. D. (2018). Exploring beyond simple demographic variables: Differences between traditional laboratory samples and crowdsourced online samples on the Big Five personality traits. *Personality and Individual Differences*, *133*, 41–46. https://doi.org/10.1016/j.paid.2017.06.023

COP24. (2018). *COP24: Key outcomes agreed at the UN climate talks in Katowice*. https://www.carbonbrief.org/cop24-key-outcomes-agreed-at-the-un-climate-talks-in-katowice

Coppock, A. (2019). Generalizing from Survey Experiments Conducted on Mechanical Turk: A Replication Approach. *Political Science Research and Methods*, *7*(3), 613–628. https://doi.org/10.1017/psrm.2018.10

Corriveau, K., & Harris, P. L. (2009). Choosing your informant: Weighing familiarity and recent accuracy. *Developmental Science*, *12*(3), 426–437. https://doi.org/10.1111/j.1467-7687.2008.00792.x

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In *The adapted mind: Evolutionary psychology and the generation of culture.* (pp. 163–288). Oxford University Press. https://doi.org/10.1080/00221325.2012.749833

Costa, D. L., & Kahn, M. E. (2013). Energy conservation “nudges” and environmentalist ideology: Evidence from a randomized residential electricity field experiment. *Journal of the European Economic Association*, *11*(3), 680–702. https://doi.org/10.1111/jeea.12011

Cotton, D. R. E., & Alcock, I. (2013). Commitment to environmental sustainability in the UK student population. *Studies in Higher Education*, *38*(10), 1457–1471. https://doi.org/10.1080/03075079.2011.627423

Craig, C. S., & McCann, J. M. (1978). Assessing Communication Effects on Energy Conservation. *Journal of Consumer Research*, *5*(2), 82–88. https://doi.org/10.1086/208718

Creswell, J. W. L., & Clark, V. L. P. (2011). *Designing and Conducting Mixed Methods Research* (2nd ed.). Sage.

Cruz, S. M. (2017). The relationships of political ideology and party affiliation with environmental concern: A meta-analysis. *Journal of Environmental Psychology*, *53*, 81–91. https://doi.org/10.1016/j.jenvp.2017.06.010

Cui, P. B., & Wu, Z. X. (2013). Impact of conformity on the evolution of cooperation in the prisoner's dilemma game. *Physica A: Statistical Mechanics and Its Applications*, *392*(6), 1500–1509. https://doi.org/10.1016/j.physa.2012.10.039

Culiberg, B., & Elgaaied-Gambier, L. (2016). Going green to fit in - understanding the impact of social norms on pro-environmental behaviour, a cross-cultural approach. *International Journal of Consumer Studies*, *40*(2), 179–185. https://doi.org/10.1111/ijcs.12241

Cuomo, M. T., Foroudi, P., Tortora, D., Hussain, S., & Melewar, T. C. (2019). Celebrity endorsement and the attitude towards luxury brands for sustainable consumption. *Sustainability (Switzerland)*, *11*(23), 1–21. https://doi.org/10.3390/su11236791

Dallinger. (2020). *Dallinger*. https://dallinger.readthedocs.io/en/latest/

Danku, Z., Wang, Z., & Szolnoki, A. (2018). Imitate or innovate: Competition of strategy updating attitudes in spatial social dilemma games. *Epl*, *121*(1), 2–9. https://doi.org/10.1209/0295-5075/121/18002

Darwin, C. (1859). On the Origin of Species by Means of Natural Selection. In *D. Appleton and Company*. https://doi.org/10.1007/s11664-006-0098-9

Davies, N. B., Krebs, J. R., & West, S. A. (2012). An Introduction to Behavioural Ecology Fourth Edition. In *Statewide Agricultural Land Use Baseline 2015*.

Davis, H., & McLeod, S. L. (2003). Why humans value sensational news. An evolutionary perspective. *Evolution and Human Behavior*, *24*(3), 208–216. https://doi.org/10.1016/S1090-5138(03)00012-6

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007). Egalitarian motives in humans. *Nature*, *446*(7137), 794–796. https://doi.org/10.1038/nature05651

Dawes, R. M. (1980). Social dilemmas. *Annu. Rev. Psychol.*, *31*, 169–193. https://doi.org/10.1080/002075900399402

Dawes, R. M., & Thaler, R. H. (1988). Anomalies: Cooperation. *Journal of Economic Perspectives*, *2*(3), 187–197. https://doi.org/10.1257/jep.2.3.187

Dawkins, R. (1976). The selfish gene. New York. *Oxford Univ. Press*. https://doi.org/10.2307/2407425

De Cremer, D., & Van Knippenberg, D. (2002). How do leaders promote cooperation? The effects of charisma and procedural fairness. *Journal of Applied Psychology*, *87*(5), 858–866. https://doi.org/10.1037/0021-9010.87.5.858

de groot, J., & Steg, L. (2008). Value Orientations to Explain Beliefs Related to Environmental Significant Behavior. *Environment and Behavior*, *40*(3), 330–354. https://doi.org/10.1177/0013916506297831

de Hooge, I. E., Zeelenberg, M., & Breugelmans, S. M. (2007). Moral sentiments and cooperation: Differential influences of shame and guilt. *Cognition and Emotion*, *21*(5), 1025–1042. https://doi.org/10.1080/02699930600980874

De Quervain, D. J. F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., & Fehr, E. (2004). The neural basis of altruistic punishment. *Science*, *305*(5688), 1254–1258. https://doi.org/10.1126/science.1100735

Dean, L. G., Vale, G. L., Laland, K. N., Flynn, E., & Kendal, R. L. (2014). Human cumulative culture: A comparative perspective. *Biological Reviews*, *89*(2), 284–301. https://doi.org/10.1111/brv.12053

Dear, K., Dutton, K., & Fox, E. (2019). Do 'watching eyes' influence antisocial behavior? A systematic review & meta-analysis. *Evolution and Human Behavior*, *40*(3), 269–280. https://doi.org/10.1016/j.evolhumbehav.2019.01.006

DEFRA. (2018). *Digest of waste and resource statistics - 2018 edition* (Issue May). www.gov.uk/government/collections/waste-and-recycling-statistics

Delmas, M. A., Fischlein, M., & Asensio, O. I. (2013). Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012. *Energy Policy*, *61*, 729–739. https://doi.org/10.1016/j.enpol.2013.05.109

Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences*, *108*(32), 13335–13340. https://doi.org/10.1073/pnas.1102131108

Diekmann, A., Przepiorka, W., & Rauhut, H. (2015). Lifting the veil of ignorance: An experiment on the contagiousness of norm violations. *Rationality and Society*, *27*(3), 309–333. https://doi.org/10.1177/1043463115593109

do Paço, A., & Laurett, R. (2018). Environmental behaviour and sustainable development. In *Leal Filho W. (eds) Encyclopedia of Sustainability in Higher Education*. Springer, Cham. https://doi.org/10.1007/978-3-319-63951-2_14-1

Doebeli, M., & Hauert, C. (2005a). Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game. *Ecology Letters*, *8*(7), 748–766. https://doi.org/10.1111/j.1461-0248.2005.00773.x

Doebeli, M., & Hauert, C. (2005b). Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game. *Ecology Letters*, *8*(7), 748–766. https://doi.org/10.1111/j.1461-0248.2005.00773.x

Doebeli, M., Hauert, C., & Killingback, T. (2006). The Evolutionary Origin of Cooperators and Defectors. *Science*, *859*(2004), 859–862. https://doi.org/10.1126/science.1101456 dos Santos, M., Rankin, D. J., & Wedekind, C. (2013). Human Cooperation Based On Punishment Reputation. *Evolution*, *67*(8), 2446–2450. https://doi.org/10.1111/evo.12108

Drouvelis, M., & Grosskopf, B. (2016). The effects of induced emotions on pro-social behaviour. *Journal of Public Economics*, *134*, 1–8. https://doi.org/10.1016/j.jpubeco.2015.12.012

D'Souza, C., Taghian, M., Lamb, P., & Peretiatko, R. (2007). Green decisions: Demographics and consumer understanding of environmental labels. *International Journal of Consumer Studies*, *31*(4), 371–376. https://doi.org/10.1111/j.1470-6431.2006.00567.x

Duffy, J., & Ochs, J. (2009). Cooperative behavior and the frequency of social interaction. *Games and Economic Behavior*, *66*(2), 785–812. https://doi.org/10.1016/j.geb.2008.07.003

Dugatkin, L. A. (2002). Cooperation in animals: An evolutionary overview. *Biology and Philosophy*, *17*(4), 459–476. https://doi.org/10.1023/A:1020573415343

Dunbar, R. I. M. (2009). The social brain hypothesis and its implications for social evolution. *Annals of Human Biology*, *36*(5), 562–572. https://doi.org/10.1080/03014460902960289

Dunbar, R. I. M., Marriott, A., & Duncan, N. D. C. (1997). Human conversational behavior. *Human Nature*, *8*(3), 231–246. https://doi.org/10.1007/BF02912493

Durham, C. A., Colby, B. G., & Longstreth, M. (1988). The Impact of State Tax Credits and Energy Prices on Adoption of Solar Energy Systems. *Land Economics*, *64*(4), 347–355. https://doi.org/10.2307/3146307

Eagle, L., Low, D., Case, P., & Vandommele, L. (2015). Attitudes of undergraduate business students toward sustainability issues. *International Journal of Sustainability in Higher Education*, *16*(5), 650–668. https://doi.org/10.1108/IJSHE-04-2014-0054

Eagly, A. H., & Riger, S. (2014). Feminism and psychology: Critiques of methods and epistemology. *American Psychologist*, *69*(7), 685–702. https://doi.org/10.1037/a0037372

Eberhart, A. K. (née M., & Naderer, G. (2017). Quantitative and qualitative insights into consumers' sustainable purchasing behaviour: a segmentation approach based on motives and heuristic cues. *Journal of Marketing Management*, *33*(13–14), 1149–1169. https://doi.org/10.1080/0267257X.2017.1371204

Efferson, C., Lalive, R., Richerson, P. J., McElreath, R., & Lubell, M. (2008). Conformists and mavericks: the empirics of frequency-dependent cultural transmission. *Evolution and Human Behavior*, *29*(1), 56–64. https://doi.org/10.1016/j.evolhumbehav.2007.08.003

Efferson, C., Vogt, S., & Fehr, E. (2020). The promise and the peril of using social influence to reverse harmful traditions. *Nature Human Behaviour*, *4*(1), 55–68. https://doi.org/10.1038/s41562-019-0768-2

Eggleston, S. J., & Turiel, E. (1983). The Development of Social Knowledge. Morality and Convention. *Cambridge University Press*. https://doi.org/10.2307/3121515

Eisner, E. W. (2004). On the art and science of qualitative research in psychology. In *Qualitative research in psychology: Expanding perspectives in methodology and design.* (pp. 17–29). American Psychological Association. https://doi.org/10.1037/10595-002

Ek, K., & Söderholm Patrik, P. (2010). The devil is in the details: Household electricity saving behavior and the role of information. *Energy Policy*, *38*(3), 1578–1587. https://doi.org/10.1016/j.enpol.2009.11.041

Emanuel, R., & Adams, J. N. (2011). College students' perceptions of campus sustainability. *International Journal of Sustainability in Higher Education*, *12*(1), 79–92. https://doi.org/10.1108/14676371111098320

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, *14*(4), 583–610. https://doi.org/10.1007/s10683-011-9283-7

Engelmann, J. M., Herrmann, E., Rapp, D. J., & Tomasello, M. (2016). Young children (sometimes) do the right thing even when their peers do not. *Cognitive Development*, *39*, 86–92. https://doi.org/10.1016/j.cogdev.2016.04.004

Eriksson, K., & Coultas, J. C. (2009). Are people really conformist-biased? An empirical test and a new mathematical model. *Journal of Evolutionary Psychology*, *7*(1), 5–21. https://doi.org/10.1556/JEP.7.2009.1.3

Eriksson, K., & Coultas, J. C. (2014). Corpses, maggots, poodles and rats: Emotional selection operating in three phases of cultural transmission of urban legends Abstract. *Journal of Cognition and Culture*, *14*, 1–26.

Erkal, N., Gangadharan, L., & Nikiforakis, N. (2011). Relative Earnings and Giving in a Real-Effort Experiment. *American Economic Review*, *101*(7), 3330–3348. https://doi.org/10.1257/AER.101.7.3330

Ethical Consumer. (2020). *Ethical Consumerism in the Pandemic*. Ethical Consumer Report. https://www.ethicalconsumer.org/sites/default/files/inline-files/Ethical Consumer Markets Report 2020.pdf

Evans, S., Gabbatiss, J., McSweeney, R., Chandrasekhar, A., Tandon, A., Viglione, G., Hausfather, Z., You, X., Joe Goodman, & Hayes, S. (2021). *COP26: Key outcomes agreed at the UN climate talks in Glasgow*. CarbonBrief Clear on Climate. https://www.carbonbrief.org/cop26-key-outcomes-agreed-at-the-un-climate-talks-in-glasgow

Faiers, A., Cook, M., & Neame, C. (2007). Towards a contemporary approach for understanding consumer behaviour in the context of domestic energy use. *Energy Policy*, *35*(8), 4381–4390. https://doi.org/10.1016/j.enpol.2007.01.003

Fawcett, T. W., Hamblin, S., & Giraldeau, L. A. (2013). Exposing the behavioral gambit: The evolution of learning and decision rules. *Behavioral Ecology*, *24*(1), 2–11. https://doi.org/10.1093/beheco/ars085

Fehr, E., & Gachter, S. (2000). Cooperation and Punishment in Public Goods Experiments. *The American Economic Review*, *90*(4), 980–994.

Fehr, E., Hoff, K., & Kshetramade, M. (2008). Spite and development. *American Economic Review*. https://doi.org/10.1257/aer.98.2.494

Fehr, E., & List, J. A. (2004). The hidden costs and returns of incentives-trust and trustworthiness among CEOs. *Journal of the European Economic Association*, *2*(5), 743–771. https://doi.org/10.1162/1542476042782297

Fehr, E., & Rockenbach, B. (2004). Human altruism: Economic, neural, and evolutionary perspectives. *Current Opinion in Neurobiology*, *14*(6), 784–790. https://doi.org/10.1016/j.conb.2004.10.007

Fehr, E., & Schurtenberger, I. (2018). Normative foundations of human cooperation. *Nature Human Behaviour*, *2*(7), 458–468. https://doi.org/10.1038/s41562-018-0385-5

Feldman, M. W., Aoki Kenichi, & Jochen, K. (1996). Individual versus social learning: evolutionary analysis in a fluctuating environment. *Anthropological Science*, *104*(4), 209–232.

FeldmanHall, O., Otto, A. R., & Phelps, E. A. (2018). Learning Moral Values: Another's Desire to Punish Enhances One's Own Punitive Behavior. *Journal of Experimental Psychology: General*, *147*(8), 1211–1224. https://doi.org/10.1037/xge0000405

Ferraro, P. J., & Vossler, C. A. (2010). The source and significance of confusion in public goods experiments. *B.E. Journal of Economic Analysis and Policy*, *10*(1). https://doi.org/10.2202/1935-1682.2006

Fischbacher, U., & Fehr, E. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, *25*(2), 63–87. https://doi.org/10.1016/S1090-5138(04)00005-4

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. https://doi.org/10.1257/aer.100.1.541

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are People Conditionally Cooperative ? Evidence from a Public Goods Experiment. *Economics Letters*, *71*(3), 397–404. https://doi.org/10.2139/ssrn.203288

Flynn, E., Turner, C., & Giraldeau, L. A. (2016). Selectivity in social and asocial learning: Investigating the prevalence, effect and development of young children's learning preferences. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(1690), 20150189. https://doi.org/10.1098/rstb.2015.0189

Foster, K. R., Wenseleers, T., & Ratnieks, F. L. W. (2001). Spite: Hamilton's unproven theory. *Annales Zoologici Fennici*, *38*(3–4), 229–238.

Fowler, J. H. (2005). Altruistic punishment and the origin of cooperation. *Proceedings of the National Academy of Sciences*, *102*(19), 7047–7049. https://doi.org/10.1073/pnas.0500938102

Fowler, J. H., & Christakis, N. A. (2010). Cooperative Behavior Cascades in Human Social Networks. *PNAS*, *107*(12), 5334–5338. https://doi.org/10.1073/pnas.0913149107

Frederickson, M. E. (2017). Mutualisms Are Not on the Verge of Breakdown. In *Trends in Ecology and Evolution* (Vol. 32, Issue 10, pp. 727–734). Elsevier Ltd. https://doi.org/10.1016/j.tree.2017.07.001

Frey, B. S., & Meier, S. (2004a). Pro-social behavior in a natural setting. *Journal of Economic Behavior and Organization*, *54*(1), 65–88. https://doi.org/10.1016/j.jebo.2003.10.001

Frey, B. S., & Meier, S. (2004b). Social comparisons and pro-social behavior: Testing “conditional cooperation” in a field experiment. *American Economic Review*, *94*(5), 1717–1722. https://doi.org/10.1257/0002828043052187

Frondel, M., & Vance, C. (2013). Heterogeneity in the Effect of Home Energy Audits: Theory and Evidence. *Environmental and Resource Economics*, *55*, 407–418. https://doi.org/10.1007/s10640-013-9632-4

Gächter, S., Nosenzo, D., Renner, E., & Sefton, M. (2012). Who makes a good leader? Cooperativeness, optimism, and leading-by-example. *Economic Inquiry*, *50*(4), 953–967. https://doi.org/10.1111/j.1465-7295.2010.00295.x

Gächter, S., & Thöni, C. (2005). Social Learning and Voluntary Cooperation Among Like-Minded People. *Journal of the European Economic Association*, *3*(2), 303–314. https://doi.org/https://doi.org/10.1162/jeea.2005.3.2-3.303

Gadagkar, R. (1993). Can animals be spiteful? *Trends in Ecology and Evolution*, *8*(7), 232–234. https://doi.org/10.1016/0169-5347(93)90196-V

Gadenne, D., Sharma, B., Kerr, D., & Smith, T. (2011). The influence of consumers’ environmental beliefs and attitudes on energy saving behaviours. *Energy Policy*, *39*(12), 7684–7694. https://doi.org/10.1016/j.enpol.2011.09.002

Galef, B. G., Dudley, K. E., & Whiskin, E. E. (2008). Social learning of food preferences in "dissatisfied" and "uncertain" Norway rats. *Animal Behaviour*, *75*(2), 631–637. https://doi.org/10.1016/j.anbehav.2007.06.024

Gambetta, D. (2005). Making Sense of Suicide Missions. In *Making Sense of Suicide Missions*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199276998.001.0001

Gamblin, B. W., Winslow, M. P., Lindsay, B., Newsom, A. W., & Kehn, A. (2017). Comparing In-Person, Sona, and Mechanical Turk Measurements of Three Prejudice-Relevant Constructs. *Current Psychology*, *36*(2), 217–224. https://doi.org/10.1007/s12144-015-9403-1

García, J., van Veelen, M., & Traulsen, A. (2014). Evil green beards: Tag recognition can also be used to withhold cooperation in structured populations. *Journal of Theoretical Biology*, *360*, 181–186. https://doi.org/10.1016/j.jtbi.2014.07.002

Gardner, A., Hardy, I. C. W., Taylor, P. D., & West, S. A. (2007). Spiteful Soldiers and Sex Ratio Conflict in Polyembryonic Parasitoid Wasps. *The American Naturalist*, *169*(4), 519–533. https://doi.org/10.1086/512107

Gardner, A., & West, S. A. (2004). Spite and the scale of competition. *Journal of Evolutionary Biology*, *17*(6), 1195–1203. https://doi.org/10.1111/j.1420-9101.2004.00775.x

Gardner, A., & West, S. A. (2010). Greenbeards. *Evolution*, *64*(1), 25–38. https://doi.org/10.1111/j.1558-5646.2009.00842.x

Ge, E., Chen, Y., Wu, J., & Mace, R. (2019). Large-scale cooperation driven by reputation, not fear of divine punishment. *Royal Society Open Science*, *6*(8). https://doi.org/10.1098/rsos.190991

Gelfand, M. J., & Harrington, J. R. (2015). The Motivational Force of Descriptive Norms: For Whom and When Are Descriptive Norms Most Predictive of Behavior? *Journal of Cross-Cultural Psychology*, *46*(10), 1273–1278. https://doi.org/10.1177/0022022115600796

Ghaffari, M., Rodrigo, P. G. K., Ekinci, Y., & Pino, G. (2021). Consumers' motivations for adopting a vegan diet: A mixed-methods approach. *International Journal of Consumer Studies*, *May*, 1–16. https://doi.org/10.1111/ijcs.12752

Gifford, R., & Nilsson, A. (2014). Personal and social factors that influence pro-environmental concern and behaviour: A review. *International Journal of Psychology*, *49*(3), 141–157. https://doi.org/10.1002/ijop.12034

Gino, F., Ayal, S., & Ariely, D. (2009). Contagion and Differentiation in Unethical Behavior The Effect of One Bad Apple on the Barrel. *Psychological Science*, *20*(3), 393–398. https://doi.org/10.1111/j.1467-9280.2009.02306.x

Gintis, H. (2000). Strong Reciprocity and Human Sociality. *Journal of Theoretical Biology*, *2*, 169–179. https://doi.org/10.1006/jtbi.2000.2111

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior*, *24*(3), 153–172. https://doi.org/10.1016/S1090-5138(02)00157-5

Giraldeau, L. A., Valone, T. J., & Templeton, J. J. (2002). Potential disadvantages of using socially acquired information. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *357*(1427), 1559–1566. https://doi.org/10.1098/rstb.2002.1065

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels. *Journal of Consumer Research*, *35*(3), 472–482. https://doi.org/10.1086/586910

Gore, J., Youk, H., & Van Oudenaarden, A. (2009). Snowdrift game dynamics and facultative cheating in yeast. *Nature*, *459*(7244), 253–256. https://doi.org/10.1038/nature07921

Gould, E. D., & Kaplan, T. R. (2011). Learning unethical practices from a co-worker: The peer effect of Jose Canseco. *Labour Economics*, *18*(3), 338–348. https://doi.org/10.1016/j.labeco.2010.10.004

Gov.uk. (2019). *Low-emission vehicles eligible for a plug-in grant*. https://www.gov.uk/plug-in-car-van-grants

Grabo, A., & van Vugt, M. (2016). Charismatic leadership and the evolution of cooperation. *Evolution and Human Behavior*, *37*(5), 399–406. https://doi.org/10.1016/j.evolhumbehav.2016.03.005

Gracia-Lazaro, C., Ferrer, A., Ruiz, G., Tarancon, A., Cuesta, J. A., Sanchez, A., & Moreno, Y. (2012). Heterogeneous networks do not promote cooperation when humans play a Prisoner's Dilemma. *Proceedings of the National Academy of Sciences*, *109*(32), 12922–12926. https://doi.org/10.1073/pnas.1206681109

Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). Young Children's Understanding of Joint Commitments. *Developmental Psychology*, *45*(5), 1430–1443. https://doi.org/10.1037/a0016122

Greene, J. C. (2007). *Mixed Methods of Social Enquiry*. Jossey-Bass.

Griffin, A. S., & West, S. A. (2002). Kin selection: fact and fiction. *Trends in Ecology & Evolution*, *17*(1), 15–21. https://doi.org/10.1016/S0169-5347(01)02355-2

Grove, M. (2018). Strong conformity requires a greater proportion of asocial learning and achieves lower fitness than a payoff-based equivalent. *Adaptive Behavior*, *26*(6), 323–333. https://doi.org/10.1177/1059712318807127

Grove, M. (2019). Evolving conformity : Conditions favouring conformist social learning over random copying. *Cognitive Systems Research*, *54*, 232–245. https://doi.org/10.1016/j.cogsys.2018.10.012

Grujić, J., & Lenaerts, T. (2020). Do people imitate when making decisions? Evidence from a spatial Prisoner's Dilemma experiment: Do people imitate when making decisions. *Royal Society Open Science*, *7*(7). https://doi.org/10.1098/rsos.200618

Gunnthorsdottir, A., Houser, D., & McCabe, K. (2007). Disposition, history and contributions in public goods experiments. *Journal of Economic Behavior and Organization*, *62*(2), 304–315. https://doi.org/10.1016/j.jebo.2005.03.008

Gurek, O., Bernd Irlenbusch, & Bettina Rockenbach. (2006). *The Competitive Advantage of Sanctioning Institutions*. *312*(April), 108–112.

Gürerk, Ö., Irlenbusch, B., & Rockenbach, B. (2014). On cooperation in open communities. *Journal of Public Economics*, *120*, 220–230. https://doi.org/10.1016/j.jpubeco.2014.10.001

Güth, W., & Kocher, M. (2013). *More than thirty years of ultimatum bargaining experiments*.

Güth, W., & Kocher, M. G. (2014). More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature. *Journal of Economic Behavior and Organization*, *108*, 396–409. https://doi.org/10.1016/j.jebo.2014.06.006

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, *3*(4), 367–388. https://doi.org/10.1016/0167-2681(82)90011-7

Hackel, L. M., & Zaki, J. (2018). Propagation of Economic Inequality Through Reciprocity and Reputation. *Psychological Science*, *29*(4), 604–613. https://doi.org/10.1177/0956797617741720

Hamilton, W. D. (1964). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, *7*(1), 1–16. https://doi.org/10.1016/0022-5193(64)90038-4

Hammersley, M. (1990). Whats wrong with ethnography? The myth of theoretical description. *Methodological Explorations*, *24*(4), 597–615.

Handgraaf, M. J. J., van Lidth de Jeude, M. A., & Appelt, K. C. (2013). Public praise vs. private pay: Effects of rewards on energy conservation in the workplace. *Ecological Economics*, *86*, 86–92. https://doi.org/10.1016/j.ecolecon.2012.11.008

Harber, K. D., & Cohen, D. J. (2005). The emotional broadcaster theory of social sharing. *Journal of Language and Social Psychology*, *24*(4), 382–400. https://doi.org/10.1177/0261927X05281426

Hardin, G. (1968). The Tragedy of the Commons. *Science*, *162*(June), 1243–1248. https://doi.org/10.1126/science.162.3859.1243

Harding, M., & Hsiaw, A. (2014). Goal setting and energy conservation. *Journal of Economic Behavior and Organization*, *107*(Part A), 209–227. https://doi.org/10.1016/j.jebo.2014.04.012

Harrell, A., Melamed, D., & Simpson, B. (2018). The strength of dynamic ties: The ability to alter some ties promotes cooperation in those that cannot be altered. *Science Advances*, *4*(12), eaau9109. https://doi.org/10.1126/sciadv.aau9109

Harrell, A., & Simpson, B. (2016). The dynamics of prosocial leadership: Power and influence in collective action groups. *Social Forces*, *94*(3), 1283–1308. https://doi.org/10.1093/sf/sov110

Harrison, G. W. (2007). House money effects in public good experiments: Comment. *Experimental Economics*, *10*(4), 429–437. https://doi.org/10.1007/s10683-006-9145-x

Hastie, R., & Kameda, T. (2005). The robust beauty of majority rules in group decisions. *Psychological Review*, *112*(2), 494–508. https://doi.org/10.1037/0033-295X.112.2.494

Hauert, C., & Doebeli, M. (2004). Spatial structure often inhibits the evolution of cooperation in the snowdrift game. In *Nature* (Vol. 428, Issue 6983, pp. 643–646). https://doi.org/10.1038/nature02360

Haun, D. B. M., Rekers, Y., & Tomasello, M. (2012). Majority-biased transmission in chimpanzees and human children, but not orangutans. *Current Biology*, *22*(8), 727–731. https://doi.org/10.1016/j.cub.2012.03.006

Hauser, M., McAuliffe, K., & Blake, P. R. (2009). Evolving the ingredients for reciprocity and spite. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1533), 3255–3266. https://doi.org/10.1098/rstb.2009.0116

Heath, C., Bell, C., & Sternberg, E. (2001). Emotional selection in memes: The case of urban legends. *Journal of Personality and Social Psychology*, *81*(6), 1028–1041. https://doi.org/10.1037//0022-3514.81.6.1028

Heath, J. (2006). The Benefits of Cooperation. *Philosophy & Public Affairs*, *34*(4), 313–351. https://doi.org/10.1111/j.1088-4963.2006.00073.x

Henrich, J. (2001). Cultural Transmission and the Diffusion of Innovations: Adoption Dynamics Indicate That Biased Cultural Transmission Is the Predominate Force in Behavioral Change. *American Anthropologist*, *103*(4), 992–1013. https://doi.org/10.1525/aa.2001.103.4.992

Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior and Organization*, *53*(1), 3–35. https://doi.org/10.1016/S0167-2681(03)00094-5

Henrich, J. (2018). The Secret of Our success. In *The Secret of Our Success*. Princeton University press. https://doi.org/https://doi.org/10.1515/9781400873296

Henrich, J., & Boyd, R. (1998). The Evolution of Conformist Transmission and the Emergence of Between-Group Differences. *Evolution and Human Behavior*, *19*(4), 215–241. https://doi.org/10.1016/S1090-5138(98)00018-X

Henrich, J., Boyd, R., & Bowles, S. (2005). "Economic Man" in Cross-Cultural Perspective: Ethnography and Experiments from 15 Small-Scale Societies. *Behavioral and Brain Sciences*, *28*(2005), 795–855.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). Cooperation , Reciprocity and Punishment in Fifteen Small-scale Societies. *America Economics Review*, *May*, 1–7. https://doi.org/10.2307/2677736

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N. S., Hill, K., Gil-White, F., Gurven, M., Marlowe, F. W., Patton, J. Q., & Tracer, D. (2005). "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, *28*(6), 795–815. https://doi.org/10.1017/S0140525X05000142

Henrich, J., & Broesch, J. (2011). On the nature of cultural transmission networks: Evidence from Fijian villages for adaptive learning biases. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1139–1148. https://doi.org/10.1098/rstb.2010.0323

Henrich, J., Chudek, M., & Boyd, R. (2015). The Big Man Mechanism: how prestige fosters cooperation and creates prosocial leaders. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1683), 20150013. https://doi.org/10.1098/rstb.2015.0013

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. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, *327*(5972), 1480–1484. https://doi.org/10.1126/science.1182238

Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior, 22*(3), 165–196. https://doi.org/10.1016/S1090-5138(00)00071-4

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33.2-3,* 61–83. https://doi.org/10.1017/S0140525X0999152X

Henrich, J., & Henrich, N. (2010). The evolution of cultural adaptations: Fijian food taboos protect against dangerous marine toxins. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1701), 3715–3724. https://doi.org/10.1098/rspb.2010.1191

Henrich, J., & Muthukrishna, M. (2021). The Origins and Psychology of Human Cooperation. *Annual Review of Psychology*, *72*(1), 1–44. https://doi.org/10.1146/annurev-psych-081920-042106

Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science, 319*(5868), 1362–1367. https://doi.org/10.1126/science.1153808

Hertel, G., Neuhof, J., Theuer, T., & Kerr, N. L. (2000). Mood effects on cooperation in small groups: Does positive mood simply lead to more cooperation? *Cognition and Emotion, 14*(4), 441–472. https://doi.org/10.1080/026999300402754

Heyes, C. (1994). Social learning in animals: Categories and mechanisms. In *Biological Reviews of the Cambridge Philosophical Society* (Vol. 69, pp. 207–231). https://doi.org/10.1111/j.1469-185x.1994.tb01506.x

Heyes, C. (2016). Who Knows? Metacognitive Social Learning Strategies. *Trends in Cognitive Sciences*, *20*(3), 204–213. https://doi.org/10.1016/j.tics.2015.12.007

Heyes, C. (2018). *Cognitive Gadgets: The Cultural Evolution of Thinking*. Harvard University Press. https://doi.org/10.1017/S0140525X18002145

Hilbe, C., Chatterjee, K., & Nowak, M. A. (2018). Partners and rivals in direct reciprocity. *Nature Human Behaviour*, *2*(7), 469–477. https://doi.org/10.1038/s41562-018-0320-9

Hilbig, B. E., Kieslich, P. J., Henninger, F., Thielmann, I., & Zettler, I. (2018). Lead Us (Not) into Temptation: Testing the Motivational Mechanisms Linking Honesty–Humility to Cooperation. *European Journal of Personality*, *32*(2), 116–127. https://doi.org/10.1002/per.2149

Hill, K. R., Wood, B. M., Baggio, J., Hurtado, A. M., & Boyd, R. T. (2014). Hunter-gatherer inter-band interaction rates: Implications for cumulative culture. *PLoS ONE*, *9*(7). https://doi.org/10.1371/journal.pone.0102806

Hillis, V., Bell, A., Brandt, J., & Brooks, J. S. (2018). Applying a cultural multilevel selection framework to the adoption of sustainable management practices in California viticulture. *Sustainability Science*, *13*(1), 71–80. https://doi.org/10.1007/s11625-017-0515-4

Hirsh, J. B. (2010). Personality and environmental concern. *Journal of Environmental Psychology*, *30*(2), 245–248. https://doi.org/10.1016/j.jenvp.2010.01.004

Hirsh, J. B. (2014). Environmental sustainability and national personality. *Journal of Environmental Psychology*, *38*, 233–240. https://doi.org/10.1016/j.jenvp.2014.02.005

Hirsh, J. B., & Dolderman, D. (2007). Personality predictors of Consumerism and Environmentalism: A preliminary study. *Personality and Individual Differences*, *43*(6), 1583–1593. https://doi.org/10.1016/j.paid.2007.04.015

Hofbauer, J., & Sigmund, K. (2003). Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, *40*(4), 479–519. https://doi.org/10.1090/S0273-0979-03-00988-1

Hoffmann, R. (2013). The experimental economics of religion. *Journal of Economic Surveys*, *27*(5), 813–845. https://doi.org/10.1111/j.1467-6419.2011.00716.x

Hood, J. C. (2006). Teaching against the Text: The Case of Qualitative Methods. *Teaching Sociology*, *34*(3), 5–48.

Horhota, M., Asman, J., Stratton, J. P., & Halfacre, A. C. (2014). Identifying behavioral barriers to campus sustainability: A multi-method approach. *International Journal of Sustainability in Higher Education*, *15*(3), 343–358. https://doi.org/10.1108/IJSHE-07-2012-0065

Horn, J. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/https://doi.org/10.1007/BF02289447

Howitt, D., & Cramer, D. (2008). Introduction to Research Methods in Psychology. In *Introduction to Research Methods in Psychology* (Second Edi). Pearson Education Limited.

Hupper, A., Chakraborty, S., & Waring, T. (2019). Using Cooperation Science to Strengthen Maine's Local Food Economy. *Maine Policy Review*, *28*(2), 23–33. https://doi.org/10.53558/goxy7199

Inglis, R. F., Roberts, P. G., Gardner, A., & Buckling, A. (2011). Spite and the scale of competition in Pseudomonas aeruginosa. *American Naturalist*, *178*(2), 276–285. https://doi.org/10.1086/660827

Iyer, S., & Killingback, T. (2020). Evolution of cooperation in social dilemmas with assortative interactions. *Games*, *11*(4), 1–31. https://doi.org/10.3390/g11040041

Jachimowicz, J. M., Hauser, O. P., O'Brien, J. D., Sherman, E., & Galinsky, A. D. (2018). The critical role of second-order normative beliefs in predicting energy conservation. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-018-0434-0

Jacquet, J., Hauert, C., Traulsen, A., & Milinski, M. (2011). Shame and honour drive cooperation. *Biology Letters*, *7*(6), 899–901. https://doi.org/10.1098/rsbl.2011.0367

Jaffe, A. B., & Stavins, R. N. (1994). The energy-efficiency gap What does it mean? *Energy Policy*, *22*(10), 804–810. https://doi.org/10.1016/0301-4215(94)90138-4

Jain, R. K., Taylor, J. E., & Peschiera, G. (2012). Assessing eco-feedback interface usage and design to drive energy efficiency in buildings. *Energy and Buildings*, *48*, 8–17. https://doi.org/10.1016/j.enbuild.2011.12.033

Jansson, F. (2015). What games support the evolution of an ingroup bias? *Journal of Theoretical Biology*, *373*, 100–110. https://doi.org/10.1016/j.jtbi.2015.03.008

Jansson, F., & Eriksson, K. (2015). Cooperation and shared beliefs about trust in the assurance game. *PLoS ONE*, *10*(12), 1–13. https://doi.org/10.1371/journal.pone.0144191

Jensen, K. (2010). Punishment and spite, the dark side of cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1553), 2635–2650. https://doi.org/10.1098/rstb.2010.0146

Jensen, K., Call, J., & Tomasello, M. (2007). Chimpanzees are vengeful but not spiteful. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(32), 13046–13050. https://doi.org/10.1073/pnas.0705555104

Jiang, L. L., Li, W. J., & Wang, Z. (2015). Multiple effect of social influence on cooperation in interdependent network games. *Scientific Reports*, *5*, 1–8. https://doi.org/10.1038/srep14657

Jiménez, Á. v., Flitton, A., & Mesoudi, A. (2021). When do people prefer dominant over prestigious political leaders? *Evolutionary Human Sciences*, *3*. https://doi.org/10.1017/ehs.2021.12

Jiménez, Á. v., & Mesoudi, A. (2019). Prestige-biased social learning: current evidence and outstanding questions. *Palgrave Communications*, *5*(1), 1–12. https://doi.org/10.1057/s41599-019-0228-7

Jiménez, Á. v., Stubbersfield, J. M., & Tehrani, J. J. (2018). An experimental investigation into the transmission of antivax attitudes using a fictional health controversy. *Social Science and Medicine*, *215*(August), 23–27. https://doi.org/10.1016/j.socscimed.2018.08.032

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of Personality: Theory and Research*, *2*(510), 102–138. https://doi.org/citeulike-article-id:3488537

John, P., James, O., Moseley, A., Ryan, M., Richardson, L., & Stoker, G. (2019). The Impact of Peer, Politician, and Celebrity Endorsements on Volunteering: A Field Experiment with English Students. *Journal of Nonprofit and Public Sector Marketing*, *31*(3), 328–346. https://doi.org/10.1080/10495142.2018.1526743

Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, *32*(5), 865–889. https://doi.org/10.1016/j.joep.2011.05.007

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research*, *1*(2), 112–133. https://doi.org/10.1002/9781119410867.ch12

Jones, P. L., Ryan, M. J., & Chittka, L. (2015). The influence of past experience with flower reward quality on social learning in bumblebees. *Animal Behaviour*, *101*, 11–18. https://doi.org/10.1016/j.anbehav.2014.12.016

Joshi, Y., & Rahman, Z. (2015). Factors Affecting Green Purchase Behaviour and Future Research Directions. *International Strategic Management Review*, *3*(1–2), 128–143. https://doi.org/10.1016/j.ism.2015.04.001

Joshi, Y., & Rahman, Z. (2019). Consumers' Sustainable Purchase Behaviour: Modeling the Impact of Psychological Factors. *Ecological Economics*, *159*(1270), 235–243. https://doi.org/10.1016/j.ecolecon.2019.01.025

Kabiri, S., Shadmanfaat, S. M., Smith, H., & Choi, J. (2020). Antisocial Behavior in Soccer Players: Using an Integrated Mediation Model of Personal Control and Social Learning Theory. *Social Science Quarterly*, *101*(3), 1090–1114. https://doi.org/10.1111/ssqu.12793

Kagel, J., & McGee, P. (2014). Personality and cooperation in finitely repeated prisoner's dilemma games. *Economics Letters*, *124*(2), 274–277. https://doi.org/10.1016/j.econlet.2014.05.034

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the Assumptions of Economics. *The Journal of Business*, *59*(4), S285–S300. https://doi.org/10.1086/296367

Kaiser, F. G., & Biel, A. (2000). Assessing General Ecological Behavior. *European Journal of Psychological Assessment*, *16*(1), 44–52. https://doi.org/10.1027//1015-5759.16.1.44

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Measurement*, *20*, 141–151. https://doi.org/https://doi.org/10.1177/001316446002000116

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31–36. https://doi.org/10.1007/BF02291575

Kandler, A., & Laland, K. N. (2009). An investigation of the relationship between innovation and cultural diversity. *Theoretical Population Biology*, *76*(1), 59–67. https://doi.org/10.1016/j.tpb.2009.04.004

Kang, S. H. K., McDermott, K. B., & Cohen, S. M. (2008). The mnemonic advantage of processing fitness-relevant information. *Memory and Cognition*, *36*(6), 1151–1156. https://doi.org/10.3758/MC.36.6.1151

Keizer, K., Lindenberg, S., & Steg, L. (2008). The spreading of disorder. *Science*, *322*(5908), 1681–1685. https://doi.org/10.1126/science.1161405

Keller, L., & Ross, K. G. (1998). Selfish genes: A green beard in the red fire ant. *Nature*, *394*(6693), 573–575. https://doi.org/10.1038/29064

Kelsey, C., Vaish, A., & Grossmann, T. (2018). Eyes, More Than Other Facial Features, Enhance Real-World Donation Behavior. *Human Nature*, *29*(4), 390–401. https://doi.org/10.1007/s12110-018-9327-1

Kendal, J. R., Rendell, L., Pike, T. W., & Laland, K. N. (2009). Nine-spined sticklebacks deploy a hill-climbing social learning strategy. *Behavioral Ecology*, *20*(2), 238–244. https://doi.org/10.1093/beheco/arp016

Kendal, R., Boogert, N. J., Rendell, L., Laland, K. N., Webster, M., & Jones, P. L. (2018). Social Learning Strategies: Bridge-Building between Fields. *Trends in Cognitive Sciences*, *22*(7), 651–665. https://doi.org/10.1016/j.tics.2018.04.003

Kendal, R., Coolen, I., van Bergen, Y., & Laland, K. N. (2005). Trade-Offs in the Adaptive Use of Social and Asocial Learning. *Advances in the Study of Behavior*, *35*(November), 333–379. https://doi.org/10.1016/S0065-3454(05)35008-X

Kendal, R., Hopper, L. M., Whiten, A., Brosnan, S. F., Lambeth, S. P., Schapiro, S. J., & Hoppitt, W. (2015). Chimpanzees copy dominant and knowledgeable individuals: Implications for cultural diversity. *Evolution and Human Behavior*, *36*(1), 65–72. https://doi.org/10.1016/j.evolhumbehav.2014.09.002

Kendal, R., Kendal, J. R., Hoppitt, W., & Laland, K. N. (2009). Identifying social learning in animal populations: A new "option-bias" method. *PLoS ONE*, *4*(8). https://doi.org/10.1371/journal.pone.0006541

Keohane, R. O., & Victor, D. G. (2016). Cooperation and discord in global climate policy. *Nature Climate Change*, *6*(6), 570–575. https://doi.org/10.1038/nclimate2937

Ketelaar, T., & Tung Au, W. (2003). The effects of feelings of guilt on the behaviour of uncooperative individuals in repeated social bargaining games: An effect-as-information interpretation of the role of emotion in social interaction. *Cognition and Emotion*, *17*(3), 429–453. https://doi.org/10.1080/02699930143000662

Keuschnigg, M., & Wolbring, T. (2015). Disorder, social capital, and norm violation: Three field experiments on the broken windows thesis. *Rationality and Society*, *27*(1), 96–126. https://doi.org/10.1177/1043463114561749

Kiers, E. T., Rousseau, R. A., West, S. A., & Denison, R. F. (2003). Host sanctions and the legume– rhizobium mutualism. *Nature*, *425*(September), 1–4. papers2://publication/uuid/3108C03F-EBB0-496C-BCEF-17FF437DCBAB

Kimbrough, E. O., & Reiss, J. P. (2012). Measuring the distribution of spitefulness. *PLoS ONE*, *7*(8), 1–8. https://doi.org/10.1371/journal.pone.0041812

Kimbrough, E. O., & Vostroknutov, A. (2016). Norms Make Preferences Social. *Journal of the European Economic Association*, *14*(3), 608–638. https://doi.org/10.1111/jeea.12152

King, A. J., Johnson, D. D. P., & van Vugt, M. (2009). The Origins and Evolution of Leadership. *Current Biology*, *19*(19), R911–R916. https://doi.org/10.1016/j.cub.2009.07.027

Kingma, S. A., Santema, P., Taborsky, M., & Komdeur, J. (2014). Group augmentation and the evolution of cooperation. In *Trends in Ecology and Evolution* (Vol. 29, Issue 8, pp. 476–484). https://doi.org/10.1016/j.tree.2014.05.013

Kline, M. A., Waring, T., & Salerno, J. (2018). Designing cultural multilevel selection research for sustainability science. *Sustainability Science*, *13*(1), 9–19. https://doi.org/10.1007/s11625-017-0509-2

Kollmuss, A., & Agyeman, J. (2002). Mind the Gap: Why do people act environmentally and what are the barriers to pro-environmental behavior? *Environmental Education Research*, *8*(3), 239–260. https://doi.org/10.1080/13504620220145401

Kollock, P. (1998). Social Dilemmas: The Anatomy of Cooperation. *Annual Review of Sociology*, *24*(1), 183–214. https://doi.org/10.1146/annurev.soc.24.1.183

Krakauer, A. H. (2005). Kin selection and cooperative courtship in wild turkeys. *Nature*, *434*(7029), 69–72. https://doi.org/10.1038/nature03325

Krasnow, M. M., & Delton, A. W. (2016). The sketch is blank: No evidence for an explanatory role for cultural group selection. In *Behavioral and Brain Sciences*. https://doi.org/10.1017/S0140525X15000163

Kruschke, J. K. (2015). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan* (Second Edi). Academic Press. https://doi.org/10.1016/c2012-0-00477-2

Kümmerli, R., Colliard, C., Fiechter, N., Petitpierre, B., Russier, F., & Keller, L. (2007). Human cooperation in social dilemmas: Comparing the Snowdrift game with the Prisoner's Dilemma. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1628), 29652970. https://doi.org/10.1098/rspb.2007.0793

Kumru, C. S., & Vesterlund, L. (2010). *The Effect of Status on Charitable Giving*. *12*(4), 709–735. https://pdfs.semanticscholar.org/ce21/dfa6c96b22241c179fc16b1868dd3c654130.pdf

Kun, Á., Boza, G., & Scheuring, I. (2006). Asynchronous snowdrift game with synergistic effect as a model of cooperation. *Behavioral Ecology*, *17*(4), 633–641. https://doi.org/10.1093/beheco/ark009

Kurland, J. A., & Gaulin, S. J. C. (2005). Cooperation and Conflict among Kin. In *The Handbook of Evolutionary Psychology*. https://doi.org/10.1002/9780470939376.ch15

Kurokawa, S., & Ihara, Y. (2017). Evolution of group-wise cooperation: Is direct reciprocity insufficient? *Journal of Theoretical Biology*, *415*(September 2016), 20–31. https://doi.org/10.1016/j.jtbi.2016.12.002

Kurzban, R., Burton-Chellew, M. N., & West, S. (2015). The Evolution of Altruism in Humans. *Annual Review of Psychology*, *66*, 575–599. https://doi.org/10.1146/annurev-psych-010814-015355

Kurzban, R., & Houser, D. (2001). Individual differences in cooperation in a circular public goods game. *European Journal of Personality*, *15*(S1), S37–S52. https://doi.org/10.1002/per.420

Kvale, S. (1995). The Social Construction of Validity. *Qualitative Inquiry*, *1*(1), 19–40.

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience, 7*(1), 54–64. https://doi.org/10.1038/nrn1825

Lakens, D. (2021). The Practical Alternative to the p Value Is the Correctly Used p Value. *Perspectives on Psychological Science, 16*(3), 639–648. https://doi.org/10.1177/1745691620958012

Laland, K. N. (2004). Social learning strategies. *Animal Learning & Behavior, 32*(1), 4–14. https://doi.org/10.3758/BF03196002

Laland, K. N., & Hoppitt, W. (2003). Do Animals Have Culture? *Evolutionary Anthropology*, *12*(3), 150–159. https://doi.org/10.1002/evan.10111

Laland, K. N., & Williams, K. (1997). Shoaling generates social learning of foraging information in guppies. *Animal Behaviour*, *53*(6), 1161–1169. https://doi.org/10.1006/anbe.1996.0318

Lamba, S. (2014). Social learning in cooperative dilemmas. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1787), 20140417. https://doi.org/10.1098/rspb.2014.0417

Lamba, S., & Mace, R. (2011). Demography and ecology drive variation in cooperation across human populations. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(35), 14426–14430. https://doi.org/10.1073/pnas.1105186108

Langbroek, J. H. M., Franklin, J. P., & Susilo, Y. O. (2016). The effect of policy incentives on electric vehicle adoption. *Energy Policy*, *94*, 94–103. https://doi.org/10.1016/j.enpol.2016.03.050

Lange, T., & Waring, T. (2018). The Science of Cooperation: Psychology provides new insights into collective Action. *Cooperative Business Journal*, 4–11. https://ncbaclusa.coop/journal/2019/summer/the-science-of-cooperation/

Lapadat, J. C. (2000). Problematizing transcription: Purpose, paradigm and quality. *International Journal of Social Research Methodology*, *3*(3), 203–219. https://doi.org/10.1080/13645570050083698

Leach, E. (1961). *Pul Eliya, a village in Ceylon: a study of land tenure and kinship*. Cambridge University Press. https://doi.org/885409

Ledyard, J. O. (1995). Public Goods: A Survey of Experimental Research. In *The Handbook of Experimental Economics*. Princeton University press.

Legare, C. H. (2017). Cumulative cultural learning: Development and diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(30), 7877–7883. https://doi.org/10.1073/pnas.1620743114

Legare, C. H., & Nielsen, M. (2015). Imitation and Innovation: The Dual Engines of Cultural Learning. *Trends in Cognitive Sciences*, *19*(11), 688–699. https://doi.org/10.1016/j.tics.2015.08.005

Lehmann, L., & Feldman, M. W. (2008). The co-evolution of culturally inherited altruistic helping and cultural transmission under random group formation. *Theoretical Population Biology*, *73*(4), 506–516. https://doi.org/10.1016/j.tpb.2008.02.004

Lehmann, L., Feldman, M. W., & Foster, K. R. (2008). Cultural Transmission Can Inhibit the Evolution of Altruistic Helping. *The American Naturalist*, *172*(1), 12–24. https://doi.org/10.1086/587851

Leigh, E. G. (2010). The evolution of mutualism. In *Journal of Evolutionary Biology* (Vol. 23, Issue 12, pp. 2507–2528). https://doi.org/10.1111/j.1420-9101.2010.02114.x

Leimgruber, K. L. (2018). The developmental emergence of direct reciprocity and its influence on prosocial behavior. *Current Opinion in Psychology*, *20*, 122–126. https://doi.org/10.1016/j.copsyc.2018.01.006

Leimgruber, K. L., Rosati, A. G., & Santos, L. R. (2016). Capuchin monkeys punish those who have more. *Evolution and Human Behavior*, *37*(3), 236–244. https://doi.org/10.1016/j.evolhumbehav.2015.12.002

Lenfesty, H. L., & Morgan, T. J. H. (2019). By Reverence, Not Fear: Prestige, Religion, and Autonomic Regulation in the Evolution of Cooperation. *Frontiers in Psychology*, *10*(December), 1–13. https://doi.org/10.3389/fpsyg.2019.02750

Leris, I., & Reader, S. M. (2016). Age and early social environment influence guppy social learning propensities. *Animal Behaviour, 120*, 11–19. https://doi.org/10.1016/j.anbehav.2016.07.012

Levine, D. S., & Strube, M. J. (2012). Environmental attitudes, knowledge, intentions and behaviors among college students. *Journal of Social Psychology*, *152*(3), 308–326. https://doi.org/10.1080/00224545.2011.604363

Levitt, S., & List, J. (2007). What do Laboratory Experiments Tell us About the Real World? *Journal of Economic Perspectives*, *21*(2), 153–174. https://doi.org/10.1257/jep.21.2.153

Lewis, C. J., Hodgkinson, E. L., & Allison, K. P. (2020). Corrosive attacks in the UK – Psychosocial perspectives and decontamination strategies. *Burns*, *46*(1), 213–218. https://doi.org/10.1016/j.burns.2019.06.003

Lewis, P. R., & Ekart, A. (2017). Social and asocial learning in collective action problems: The rise and fall of socially-beneficial behaviour. *Proceedings - 2017 IEEE 2nd International Workshops on Foundations and Applications of Self* Systems, FAS*W 2017*, 91–96. https://doi.org/10.1109/FAS-W.2017.126

Li, X., Molleman, L., & van Dolder, D. (2021). Do descriptive social norms drive peer punishment? Conditional punishment strategies and their impact on cooperation. *Evolution and Human Behavior, 42*(5), 469–479. https://doi.org/10.1016/j.evolhumbehav.2021.04.002

Lin, M., Lucas, H. C., & Shmueli, G. (2013). Too big to fail: Large samples and the p-value problem. *Information Systems Research*, *24*(4), 906–917. https://doi.org/10.1287/isre.2013.0480

List, J. A. (2006). The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions. *Journal of Political Economy*, *114*(1), 1–37. https://doi.org/10.1086/498587

Little, A. C., Caldwell, C. A., Jones, B. C., & DeBruine, L. M. (2015). Observer age and the social transmission of attractiveness in humans: Younger women are more influenced by the choices of popular others than older women. *British Journal of Psychology*, *106*(3), 397–413. https://doi.org/10.1111/bjop.12098

Lönnqvist, J. E., Verkasalo, M., & Walkowitz, G. (2011). It pays to pay - Big Five personality influences on co-operative behaviour in an incentivized and hypothetical prisoner's dilemma game. *Personality and Individual Differences*, *50*(2), 300–304. https://doi.org/10.1016/j.paid.2010.10.009

Mace, R., & Silva, A. S. (2016). The role of cultural group selection inexplaining human cooperation is a hard case to prove. In *Behavioral and Brain Sciences* (Vol. 39, p. e45). https://doi.org/10.1017/S0140525X15000187

Madsen, E. A., Tunney, R. J., Fieldman, G., Plotkin, H. C., Dunbar, R. I. M., Richardson, J. M., & McFarland, D. (2007). Kinship and altruism: A cross-cultural experimental study. *British Journal of Psychology*, *98*(2), 339–359. https://doi.org/10.1348/000712606X129213

Mahajan, N., & Wynn, K. (2012). Origins of "Us" versus "Them": Prelinguistic infants prefer similar others. *Cognition*, *124*(2), 227–233. https://doi.org/10.1016/j.cognition.2012.05.003

Mahmood, L., Abrams, D., Meleady, R., Hopthrow, T., Lalot, F., Swift, H., & van de Vyver, J. (2019). Intentions, efficacy, and norms: The impact of different self-regulatory cues on reducing engine idling at long wait stops. *Journal of Environmental Psychology*, *66*(April). https://doi.org/10.1016/j.jenvp.2019.101368

Majolo, B., & Maréchal, L. (2017). Between-group competition elicits within-group cooperation in children. *Scientific Reports*, *7*(November 2016), 1–9. https://doi.org/10.1038/srep43277

Malete, L., Chow, G. M., & Feltz, D. L. (2013). Influence of coaching efficacy and coaching competency on athlete-level moral variables in Botswana youth soccer. *Journal of Applied Social Psychology*, *43*(10), 2107–2119. https://doi.org/10.1111/jasp.12164

Malodia, S., & Bhatt, A. S. (2019). Why Should I Switch Off: Understanding the Barriers to Sustainable Consumption? *Vision*, *23*(2), 134–143. https://doi.org/10.1177/0972262919840197

Marczyk, J. (2017). Human punishment is not primarily motivated by inequality. *PLoS ONE*, *12*(2), 1–11. https://doi.org/10.1371/journal.pone.0171298

Marlowe, F. W., Berbesque, J. C., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., Ensminger, J., Gurven, M., Gwako, E., Henrich, J., Henrich, N., Lesorogol, C., McElreath, R., & Tracer, D. (2008). More “altruistic” punishment in larger societies. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1634), 587–592. https://doi.org/10.1098/rspb.2007.1517

Martin, J. S., Koski, S. E., Bugnyar, T., Jaeggi, A. v., & Massen, J. J. M. (2021). Prosociality, social tolerance and partner choice facilitate mutually beneficial cooperation in common marmosets, Callithrix jacchus. *Animal Behaviour*, *173*, 115–136. https://doi.org/10.1016/j.anbehav.2020.12.016

Matsuura, K., Fujimoto, M., Goka, K., & Nishida, T. (2002). Cooperative colony foundation by termite female pairs: Altruism for survivorship in incipient colonies. *Animal Behaviour*, *64*(2), 167–173. https://doi.org/10.1006/anbe.2002.3062

Maynard Smith, J., & Price, G. R. (1973). The logic of animal conflict. *Nature*, *246*, 15–18. https://doi.org/10.1038/246015a0

Mcandrew, F. T., Bell, K. E., & Contitta, M. G. (2007). Who Do We Tell and Whom Do We Tell On ? Gossip as a Strategy for Status Enhancement. *Journal of Applied Social Psychology*, *37*, 1562–1577.

McElreath, R. (2016). Statistical Rethinking. In *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. https://doi.org/10.3102/1076998616659752

McElreath, R. (2020). Statistical Rethinking. In *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. CRC Press. https://doi.org/10.1201/9780429029608

McElreath, R., Bell, A. v., Efferson, C., Lubell, M., Richerson, P. J., & Waring, T. (2008). Beyond existence and aiming outside the laboratory: Estimating frequency-dependent and pay-off-biased social learning strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1509), 3515–3528. https://doi.org/10.1098/rstb.2008.0131

McElreath, R., & Henrich, J. (2003). The Evolution of Cultural Evolution. *Evolutionary Anthropology*, *12*(3), 123–135. https://doi.org/10.1002/evan.10110

McGuigan, N., & Cubillo, M. (2013). Information transmission in young children: When social information is more important than nonsocial information. *Journal of Genetic Psychology*, *174*(6), 605–619. https://doi.org/10.1080/00221325.2012.749833

Mcnamara, J. M., Barta, Z., Fromhage, L., & Houston, A. I. (2008). The coevolution of choosiness and cooperation. *Nature*, *451*(January), 189–192. https://doi.org/10.1038/nature06455

McNamara, K. (2011). The paparazzi industry and new media: The evolving production and consumption of celebrity news and gossip websites. *International Journal of Cultural Studies*, *14*(5), 515–530. https://doi.org/10.1177/1367877910394567

Merle, M., Reese, G., & Drews, S. (2019). #Globalcitizen: An explorative Twitter analysis of global identity and sustainability communication. *Sustainability (Switzerland)*, *11*(12), 1–10. https://doi.org/10.3390/SU11123472

Mesoudi, A. (2009). The cultural dynamics of copycat suicide. *PLoS ONE*, *4*(9), e7252. https://doi.org/10.1371/journal.pone.0007252

Mesoudi, A. (2011). An experimental comparison of human social learning strategies: Payoff-biased social learning is adaptive but underused. *Evolution and Human Behavior*, *32*(5), 334–342. https://doi.org/10.1016/j.evolhumbehav.2010.12.001

Mesoudi, A. (2016). Cultural Evolution: A Review Of Theory, Findings and Controversies. *Evolutionary Biology*, *20*(1), 319–387. https://doi.org/10.1007/978-1-4615-6983-1

Mesoudi, A. (2020). Cultural evolution of football tactics: Strategic social learning in managers' choice of formation. *Evolutionary Human Sciences*, *2*, 1–14. https://doi.org/10.1017/ehs.2020.27

Mesoudi, A., Chang, L., Dall, S. R. X., & Thornton, A. (2016). The Evolution of Individual and Cultural Variation in Social Learning. *Trends in Ecology and Evolution*, *31*(3), 215–225. https://doi.org/10.1016/j.tree.2015.12.012

Mesoudi, A., Chang, L., Murray, K., & Lu, H. J. (2014). Higher frequency of social learning in China than in the West shows cultural variation in the dynamics of cultural evolution. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1798), 20142209. https://doi.org/10.1098/rspb.2014.2209

Mesoudi, A., & Lycett, S. J. (2009). Random copying, frequency-dependent copying and culture change. *Evolution and Human Behavior*, *30*(1), 41–48. https://doi.org/10.1016/j.evolhumbehav.2008.07.005

Mesoudi, A., & Thornton, A. (2018). What is cumulative cultural evolution? *Proceedings of the Royal Society B: Biological Sciences*, *285*(1880), 20180712. https://doi.org/10.1098/rspb.2018.0712

Mesoudi, A., & Whiten, A. (2008). The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1509), 3489–3501. https://doi.org/10.1098/rstb.2008.0129

Mesoudi, A., Whiten, A., & Dunbar, R. (2006). A bias for social information in human cultural transmission. *British Journal of Psychology*, *97*(3), 405–431. https://doi.org/10.1348/000712605X85871

Midden, C. J. H., Meter, J. F., Weenig, M. H., & Zieverink, H. J. A. (1983). Using feedback, reinforcement and information to reduce energy consumption in households: A field-experiment. *Journal of Economic Psychology, 3*, 65–86. https://doi.org/10.1016/0167-4870(83)90058-2

Milinski, M., Semmann, D., Bakker, T. C. M., & Krambeck, H. J. (2001). Cooperation through indirect reciprocity: Image scoring or standing strategy? *Proceedings of the Royal Society B: Biological Sciences*, *268*(1484), 2495–2501. https://doi.org/10.1098/rspb.2001.1809

Milinski, M., Semmann, D., Krambeck, H.-J., & Marotzke, J. (2006). Stabilizing the Earth's climate is not a losing game: Supporting evidence from public goods experiments. *Proceedings of the National Academy of Sciences*, *103*(11), 3994–3998. https://doi.org/10.1073/pnas.0504902103

Minton, A. P., & Rose, R. L. (1997). The Effects of Environmental Concern on Environmentally Friendly Consumer Behavior: An Exploratory Study. *Journal of Business Research*, *40*(1), 37–48. https://doi.org/10.1016/S0148-2963(96)00209-3

Molesti, S., & Majolo, B. (2017). Evidence of direct reciprocity, but not of indirect and generalized reciprocity, in the grooming exchanges of wild Barbary macaques (Macaca sylvanus). *American Journal of Primatology*, *79*(9), 1–12. https://doi.org/10.1002/ajp.22679

Molleman, L., Pen, I., & Weissing, F. J. (2013). Effects of Conformism on the Cultural Evolution of Social Behaviour. *PLoS ONE*, *8*(7), e68153. https://doi.org/10.1371/journal.pone.0068153

Molleman, L., Quiñones, A. E., & Weissing, F. J. (2013). Cultural evolution of cooperation: The interplay between forms of social learning and group selection. *Evolution and Human Behavior*, *34*(5), 342–349. https://doi.org/10.1016/j.evolhumbehav.2013.06.001

Molleman, L., van den Berg, P., & Weissing, F. J. (2014). Consistent individual differences in human social learning strategies. *Nature Communications*, *5*, 1–9. https://doi.org/10.1038/ncomms4570

Morgan, T. J. H., Acerbi, A., & van Leeuwen, E. J. C. (2019). Copy-the-majority of instances or individuals? Two approaches to the majority and their consequences for conformist decision-making. *PLoS ONE*, *14*(1), 1–17. https://doi.org/10.1371/JOURNAL.PONE.0210748

Morgan, T. J. H., & Laland, K. N. (2012). The biological bases of conformity. *Frontiers in Neuroscience*, *6*(JUN), 1–7. https://doi.org/10.3389/fnins.2012.00087

Morgan, T. J. H., Laland, K. N., & Harris, P. L. (2015). The development of adaptive conformity in young children: Effects of uncertainty and consensus. *Developmental Science*, *18*(4), 511–524. https://doi.org/10.1111/desc.12231

Morgan, T. J. H., Rendell, L. E., Ehn, M., Hoppitt, W., & Laland, K. N. (2012). The evolutionary basis of human social learning. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1729), 653–662. https://doi.org/10.1098/rspb.2011.1172

Morgan, T. J. H., & Thompson, B. (2020). Biased transformation erases traditions sustained by conformist transmission: Biases erase cultural traditions. *Biology Letters*, *16*(11), 20200660. https://doi.org/10.1098/rsbl.2020.0660rsbl20200660

Morin, O. (2014). Is Cooperation a Maladaptive By-product of Social Learning? The Docility Hypothesis Reconsidered. *Biological Theory*, *9*(3), 286–295. https://doi.org/10.1007/s13752-014-0181-z

Mtutu, P., & Thondhlana, G. (2016). Encouraging pro-environmental behaviour: Energy use and recycling at Rhodes University, South Africa. *Habitat International*, *53*, 142–150. https://doi.org/10.1016/j.habitatint.2015.11.031

Mujcic, R., & Leibbrandt, A. (2018). Indirect Reciprocity and Prosocial Behaviour: Evidence from a Natural Field Experiment. *Economic Journal*, *128*(611), 1683–1699. https://doi.org/10.1111/ecoj.12474

Mukherjee, D. (2009). Impact of Celebrity Endorsements on Brand Image. *Social Science Research Network*, *August*, 1–35. https://doi.org/10.1108/17473610910986026

Murphy, L. (2014). The influence of energy audits on the energy efficiency investments of private owner-occupied households in the Netherlands. *Energy Policy*, *65*, 398–407. https://doi.org/10.1016/j.enpol.2013.10.016

Muteswa, R. (2016). Qualities of a Good Leaders and the Benefits of Good Leadership to an Organization: A Conceptual Study. *European Journal of Business and Management*, *8*(24), 136.

Muthukrishna, M., Morgan, T. J. H., & Henrich, J. (2016). The when and who of social learning and conformist transmission. *Evolution and Human Behavior*, *37*(1), 10–20. https://doi.org/10.1016/j.evolhumbehav.2015.05.004

N. Bhuian, S., A. Amyx, D., & M. Shamma, H. (2014). An extension of consumer environmental behavior research among expatriates. *International Journal of Commerce and Management*, *24*(1), 63–84. https://doi.org/10.1108/IJCoMA-01-2012-0005

Nairne, J. S., & Pandeirada, J. N. S. (2008). Adaptive memory: Is survival processing special? *Journal of Memory and Language*, *59*(3), 377–385. https://doi.org/10.1016/j.jml.2008.06.001

Neilands, P., Hassall, R., Derks, F., Bastos, A. P. M., & Taylor, A. H. (2020). Watching eyes do not stop dogs stealing food: evidence against a general risk-aversion hypothesis for the watching-eye effect. *Scientific Reports*, *10*(1), 1–8. https://doi.org/10.1038/s41598-020-58210-4

Nettle, D., Colléony, A., & Cockerill, M. (2011). Variation in cooperative behaviour within a single city. *PLoS ONE*, *6*(10), e26922. https://doi.org/10.1371/journal.pone.0026922

Neugebauer, T., Perote, J., Schmidt, U., & Loos, M. (2009). Selfish-biased conditional cooperation: On the decline of contributions in repeated public goods experiments. *Journal of Economic Psychology*, *30*, 52–60. https://doi.org/10.1016/j.joep.2008.04.005

Neumann, R. (2019). The framing of charitable giving: A field experiment at bottle refund machines in Germany. *Rationality and Society*, *31*(1), 98–126. https://doi.org/10.1177/1043463118820894

Nguyen, H. V., Nguyen, C. H., & Hoang, T. T. B. (2019). Green consumption: Closing the intention-behavior gap. *Sustainable Development*, *27*(1), 118–129. https://doi.org/10.1002/sd.1875

Nguyen, T. T. H., Yang, Z., Nguyen, N., Johnson, L. W., & Cao, T. K. (2019). Greenwash and green purchase intention: The mediating role of green skepticism. *Sustainability*, *11*(9), 1–16. https://doi.org/10.3390/su11092653

Nichols, A. L., & Maner, J. K. (2008). The good-subject effect: Investigating participant demand characteristics. *Journal of General Psychology*, *135*(2), 151–166. https://doi.org/10.3200/GENP.135.2.151-166

Nisbett, R. E., & Miyamoto, Y. (2005). The influence of culture: Holistic versus analytic perception. *Trends in Cognitive Sciences*, *9*(10), 467–473. https://doi.org/10.1016/j.tics.2005.08.004

Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2008). Normative social influence is underdetected. *Personality and Social Psychology Bulletin*, *34*(7), 913–923. https://doi.org/10.1177/0146167208316691

Norbutas, L., & Corten, R. (2018). Sustainability of generalized exchange in the sharing economy: The case of the “freecycling” facebook groups. *International Journal of the Commons*, *12*(1), 111–133. https://doi.org/10.18352/ijc.789

Norenzayan, A., & Shariff, A. F. (2008). The origin and evolution of religious prosociality. In *Science*. https://doi.org/10.1126/science.1158757

Nowak, M. A. (2006). Five Rules for the Evolution of Cooperation. *Science*. https://doi.org/10.1097/00007890-197012000-00005

Nowak, M. A., & May, R. M. (1993). The spatial Dilemmas of evolution. *International Journal of Bifurcation and Chaos*, *3*(1), 35–78. https://doi.org/10.1142/S0218127493000040

Nowak, M. A., Page, K. M., & Sigmund, K. (2000). Fairness versus reason in the Ultimatum Game. *Science*, *289*(5485), 1773–1775. https://doi.org/10.1126/science.289.5485.1773

Nowak, M. A., Sasaki, A., Taylor, C., & Fudenherg, D. (2004). Emergence of cooperation and evolutionary stability in finite populations. *Nature*. https://doi.org/10.1038/nature02414

Nowak, M. A., & Sigmund, K. (1992). Tit for tat in heterogeneous populations. *Nature*, *355*(6357), 250–253. https://doi.org/10.1038/355250a0

Nowak, M. A., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. In *Nature* (Vol. 364, Issue 6432, pp. 56–58). https://doi.org/10.1038/364056a0

Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, *393*, 573–577. https://doi.org/https://doi.org/10.1038/31225

Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, *437*(7063), 1291–1298. https://doi.org/10.1038/nature04131

Nowak, M. A., Tarnita, C. E., & Wilson, E. O. (2010). The evolution of eusociality. In *Nature*. https://doi.org/10.1038/nature09205

Offord, M., Gill, R., & Kendal, J. (2016). Leadership between decks: A synthesis and development of engagement and resistance theories of leadership based on evidence from practice in Royal Navy warships. *Leadership and Organization Development Journal*, *37*(2), 289–304. https://doi.org/10.1108/LODJ-07-2014-0119

Offord, M., Gill, R., & Kendal, J. (2019). The effects of prestige on collective performance and information flow in a strictly hierarchical institution. *Palgrave Communications*, *5*(1), 1–11. https://doi.org/10.1057/s41599-018-0211-8

O'Garra, T., & Alfredo, K. A. (2019). Communication, observability and cooperation: A field experiment on collective water management in India. *Water Resources and Economics*. https://doi.org/10.1016/j.wre.2018.12.002

O'Gorman, R., Wilson, D. S., & Miller, R. R. (2008). An evolved cognitive bias for social norms. *Evolution and Human Behavior*. https://doi.org/10.1016/j.evolhumbehav.2007.07.002

Ohtsuki, H., Hauert, C., Lieberman, E., & Nowak, M. A. (2006). A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, *441*(7092), 502–505. https://doi.org/10.1038/nature04605

Ohtsuki, H., & Iwasa, Y. (2004). How should we define goodness? - Reputation dynamics in indirect reciprocity. *Journal of Theoretical Biology*. https://doi.org/10.1016/j.jtbi.2004.06.005

Ohtsuki, H., & Iwasa, Y. (2006). The leading eight: Social norms that can maintain cooperation by indirect reciprocity. *Journal of Theoretical Biology*, *239*(4), 435–444. https://doi.org/10.1016/j.jtbi.2005.08.008

Olsson, P., Folke, C., & Hahn, T. (2004). Social-ecological transformation for ecosystem management: The development of adaptive co-management of a wetland landscape in southern Sweden. *Ecology and Society*, *9*(4), 1–26. https://doi.org/10.5751/ES-00683-090402

Onu, D., & Oats, L. (2014). Social Norms and Tax Compliance. In *Tax Administration Research Center*.

Ostrom, E. (1990). Governing the commons: The evolution of institutions for collective action. In *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press. https://doi.org/10.1017/CBO9781316423936

Ostrom, E. (2000). Collective Action and the Evolution of Social Norms. *Journal of Economic Perspectives*, *14*(3), 137–158. https://doi.org/10.1080/19390459.2014.935173

Otake, R., & Dobata, S. (2018). Copy if dissatisfied, innovate if not: contrasting egg-laying decision making in an insect. *Animal Cognition*, *21*(6), 805–812. https://doi.org/10.1007/s10071-018-1212-0

Otter.ai. (2021). *Otter.ai*. https://otter.ai/

Pedersen, E. J., Kurzban, R., & McCullough, M. E. (2013). Do humans really punish altruistically? a closer look. *Proceedings of the Royal Society B: Biological Sciences*, *280*, 20122723. https://doi.org/10.1098/rspb.2012.2723

Pedersen, E. J., McAuliffe, W. H. B., & McCullough, M. E. (2018). The unresponsive avenger: More evidence that disinterested third parties do not punish altruistically. *Journal of Experimental Psychology: General*, *147*(4), 514–544. https://doi.org/10.1037/xge0000410

Perc, M., & Szolnoki, A. (2010). Coevolutionary games-A mini review. *BioSystems*, *99*(2), 109–125. https://doi.org/10.1016/j.biosystems.2009.10.003

Perry, S. (2011). Social traditions and social learning in capuchin monkeys (Cebus). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 988–996. https://doi.org/10.1098/rstb.2010.0317

Peschiera, G., & Taylor, J. E. (2012). The impact of peer network position on electricity consumption in building occupant networks utilizing energy feedback systems. *Energy and Buildings*, *49*, 584–590. https://doi.org/10.1016/j.enbuild.2012.03.011

Peters, K., Kashima, Y., & Clark, A. (2009). Talking about others: Emotionality and the dissemination of social information. *European Journal of Social Psychology*, *39*(2), 207–222. https://doi.org/10.1002/ejsp.523

Pfeiffer, T., Rutte, C., Killingback, T., Taborsky, M., & Bonhoeffer, S. (2005). Evolution of cooperation by generalized reciprocity. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1568), 1115–1120. https://doi.org/10.1098/rspb.2004.2988

Phua, J., Jin, S. V., & Kim, J. (2020). The roles of celebrity endorsers' and consumers' vegan identity in marketing communication about veganism. *Journal of Marketing Communications*, *26*(8), 813–835. https://doi.org/10.1080/13527266.2019.1590854

Pierotti, R. (1980). Spite and Altruism in Gulls. *The American Naturalist*, *115*(2), 290–300. https://doi.org/10.1086/283561

Pisor, A. C., Gervais, M. M., Purzycki, B. G., & Ross, C. T. (2020). Preferences and constraints: the value of economic games for studying human behaviour. *Royal Society Open Science*, *7*(6), 192090. https://doi.org/10.1098/rsos.192090

Player, A., Abrams, D., van de Vyver, J., Meleady, R., Leite, A. C., Randsley de Moura, G., & Hopthrow, T. (2018). "We aren't idlers": Using subjective group dynamics to promote prosocial driver behavior at long-wait stops. *Journal of Applied Social Psychology*, *48*(11), 643–648. https://doi.org/10.1111/jasp.12554

Powers, S. T., Van Schaik, C. P., & Lehmann, L. (2016). How institutions shaped the last major evolutionary transition to large-scale human societies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1687). https://doi.org/10.1098/rstb.2015.0098

Prediger, S., Vollan, B., & Herrmann, B. (2014). Resource scarcity and antisocial behavior. *Journal of Public Economics*, *119*, 1–9. https://doi.org/10.1016/j.jpubeco.2014.07.007

Priestley, M., & Mesoudi, A. (2015). Do online voting patterns reflect evolved features of human cognition? An exploratory empirical investigation. *PLoS ONE*, *10*(6), 1–15. https://doi.org/10.1371/journal.pone.0129703

Puurtinen, M., & Mappes, T. (2009a). Between-group competition and human cooperation. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1655), 355–360. https://doi.org/10.1098/rspb.2008.1060

Puurtinen, M., & Mappes, T. (2009b). Between-group competition and human cooperation. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1655), 355–360. https://doi.org/10.1098/rspb.2008.1060

Quan, J., Yang, X., & Wang, X. (2018). Spatial public goods game with continuous contributions based on Particle Swarm Optimization learning and the evolution of cooperation. *Physica A: Statistical Mechanics and Its Applications*, *505*, 973–983. https://doi.org/10.1016/j.physa.2018.04.003

Queller, D. C., Ponte, E., Bozzaro, S., & Strassmann, J. E. (2003). Single-Gene Greenbeard Effects in the Social Amoeba Dictyostelium discoideum. *Science*, *299*(5603), 105–106.

R Core Team. (2021). A language and environment for statistical computing. In *R Foundation for Statistical Computing*. https://doi.org/10.1007/978-3-540-74686-7

Raihani, N. J., & Bshary, R. (2015). Why humans might help strangers. *Frontiers in Behavioral Neuroscience*, *9*(February), 1–11. https://doi.org/10.3389/fnbeh.2015.00039

Raihani, N. J., Thornton, A., & Bshary, R. (2012). Punishment and cooperation in nature. In *Trends in Ecology and Evolution* (Vol. 27, Issue 5, pp. 288–295). Elsevier Ltd. https://doi.org/10.1016/j.tree.2011.12.004

Ramsey, G. (2013). Culture in humans and other animals. *Biology and Philosophy*, *28*(3), 457–479. https://doi.org/10.1007/s10539-012-9347-x

Rand, D. G. (2012). The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments. *Journal of Theoretical Biology*, *299*, 172–179. https://doi.org/10.1016/j.jtbi.2011.03.004

Rand, D. G., & Nowak, M. A. (2011). The evolution of anti-social punishment in optional public goods games. *Nature Communications*, *49*(18), 1841–1850. https://doi.org/10.1016/j.jacc.2007.01.076.White

Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in Cognitive Sciences*, *17*(8), 413–425. https://doi.org/10.1016/j.tics.2013.06.003

Rand, D. G., Nowak, M. A., Fowler, J. H., & Christakis, N. A. (2014). Static network structure can stabilize human cooperation. *Proceedings of the National Academy of Sciences*, *111*(48), 17093–17098. https://doi.org/10.1073/pnas.1400406111

Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., & Greene, J. D. (2014). Social heuristics shape intuitive cooperation. *Nature Communications*, *5*. https://doi.org/10.1038/ncomms4677

Rapoport, A. (1974). Prisoner's Dilemma - Recollections and observations. In *Game Theory as a theory of a conflict resolution* (pp. 17–34). Springer, Dordecht.

Rawlings, B., Flynn, E., & Kendal, R. (2017). To Copy or To Innovate? The Role of Personality and Social Networks in Children's Learning Strategies. *Child Development Perspectives*, *11*(1), 39–44. https://doi.org/10.1111/cdep.12206

Rebele, R. W., Koval, P., & Smillie, L. D. (2021). Personality-informed intervention design: Examining how trait regulation can inform efforts to change behavior. *European Journal of Personality*, *35*(4), 623–645. https://doi.org/10.1177/08902070211016251

Reed, L. I., Best, C. K., & Hooley, J. M. (2018). Cooperation with characters: How a partner's personality disorder decreases cooperation in two economic games. *Personality and Individual Differences*, *126*(January), 33–37. https://doi.org/10.1016/j.paid.2018.01.008

Rendell, L., Fogarty, L., Hoppitt, W. J. E., Morgan, T. J. H., Webster, M. M., & Laland, K. N. (2011). Cognitive culture: Theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences*, *15*(2), 68–76. https://doi.org/10.1016/j.tics.2010.12.002

Reysen, M. B., Talbert, N. G., Dominko, M., Jones, A. N., & Kelley, M. R. (2011). The effects of collaboration on recall of social information. *British Journal of Psychology*, *102*(3), 646–661. https://doi.org/10.1111/j.2044-8295.2011.02035.x

Richerson, P. J., Baldini, R., Bell, A. v., Demps, K., Frost, K., Hillis, V., Mathew, S., Newton, E. K., Naar, N., Newson, L., Ross, C., Smaldino, P. E., Waring, T. M., & Zefferman, M. (2016). Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioral and Brain Sciences*, *39*(2016), e30. https://doi.org/10.1017/S0140525X1400106X

Rieucau, G., & Giraldeau, L. A. (2011). Exploring the costs and benefits of social information use: An appraisal of current experimental evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 949–957. https://doi.org/10.1098/rstb.2010.0325

Ripple, W. J., Wolf, C., & Newsome, T. M. (2021). World Scientists ' Warning of a Climate Emergency. *Bioscience Magazine*, *71*(9), 894–898.

Ritchie, H., & Roser, M. (2020). *CO_2 and Greenhouse Gas Emissions*. Our World in Data. https://ourworldindata.org/emissions-by-sector#agriculture-forestry-and-land-use-18-4%0Ahttps://ourworldindata.org/emissions-by-sector#

Roberts, G. (2005). Cooperation through interdependence. *Animal Behaviour*, *70*(4), 901–908.

Roberts, G. (2008). Evolution of direct and indirect reciprocity. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1631), 173–179. https://doi.org/10.1098/rspb.2007.1134

Roberts, G., & Sherratt, T. N. (1998). Development of cooperative relationships through increasing investment. *Nature*, *394*(9), 175–179. https://doi.org/10.1038/28160

Roberts, Gilbert. (1998). Competitive altruism: From reciprocity to the handicap principle. *Proceedings of the Royal Society B: Biological Sciences*, *265*(1394), 427–431. https://doi.org/10.1098/rspb.1998.0312

Robertson, J. L., & Barling, J. (2013). Greening organizations through leaders' influence on employees' pro-environmental behaviors. *Journal of Organizational Behavior*, *34*(2), 176–194. https://doi.org/10.1002/job.1820

Roca, C. P., Cuesta, J. A., & Sánchez, A. (2009). Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics. *Physics of Life Reviews*, *6*(4), 208–249. https://doi.org/10.1016/j.plrev.2009.08.001

Rogers, A. R. (1988). Does Biology Constrain Culture. *American Anthropologist*, *90*(4), 819–831. https://doi.org/10.1525/aa.1988.90.4.02a00030

Romano, A., & Balliet, D. (2017a). Reciprocity Outperforms Conformity to Promote Cooperation. *Psychological Science*, *28*(10), 1490–1502. https://doi.org/10.1177/0956797617714828

Ross, S. R., Rausch, M. K., & Canada, K. E. (2003). Competition and Cooperation in the Five-Factor Model: Individual Differences in Achievement Orientation. *The Journal of Psychology*. https://doi.org/10.1080/00223980309600617

Rotella, A., Sparks, A. M., Mishra, S., & Barclay, P. (2021). No effect of 'watching eyes': An attempted replication and extension investigating individual differences. *PLoS ONE*, *16*(10 October), 1–17. https://doi.org/10.1371/journal.pone.0255531

Rothschild, C. G. (2009). The evolution of reciprocity in sizable human groups. *Journal of Theoretical Biology*, *257*(4), 609–617. https://doi.org/10.1016/j.jtbi.2009.01.004

Sabherwal, A., Ballew, M. T., van der Linden, S., Gustafson, A., Goldberg, M. H., Maibach, E. W., Kotcher, J. E., Swim, J. K., Rosenthal, S. A., & Leiserowitz, A. (2021). The Greta

Thunberg Effect: Familiarity with Greta Thunberg predicts intentions to engage in climate activism in the United States. *Journal of Applied Social Psychology*, *51*(4), 321–333. https://doi.org/10.1111/jasp.12737

Sadrieh, A., & Schröder, M. (2017). Acts of helping and harming. *Economics Letters*, *153*, 77–79. https://doi.org/10.1016/j.econlet.2017.01.019

Salazar, H. A., Oerlemans, L., & van Stroe-Biezen, S. (2013). Social influence on sustainable consumption: Evidence from a behavioural experiment. *International Journal of Consumer Studies*, *37*(2), 172–180. https://doi.org/10.1111/j.1470-6431.2012.01110.x

Sapolsky, R. M., & Share, L. J. (2004). A pacific culture among wild baboons: Its emergence and transmission. *PLoS Biology*, *2*(4), e106. https://doi.org/10.1371/journal.pbio.0020106

Sarin, S., & Dukas, R. (2009). Social learning about egg-laying substrates in fruitflies. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1677), 4323–4328. https://doi.org/10.1098/rspb.2009.1294

Sasaki, T., & Okada, I. (2015). Cheating is evolutionarily assimilated with cooperation in the continuous snowdrift game. *BioSystems*, *131*, 51–59. https://doi.org/10.1016/j.biosystems.2015.04.002

Schlag, K. H. (1996a). Why Imitate, and if so, How? A Bounded Rational Approach to Multi-Armed Bandits. *Discussion Paper Serie B*. https://doi.org/ftp://ftp.wipol.uni-bonn.de/pub/RePEc/bon/bonsfb/bonsfb361.pdf

Schlag, K. H. (1996b). Why Imitate, and if so, How? A Bounded Rational Approach to Multi-Armed Bandits. *Discussion Paper Serie B*. https://doi.org/ftp://ftp.wipol.uni-bonn.de/pub/RePEc/bon/bonsfb/bonsfb361.pdf

Schlag, K. H. (1998). Why Imitate, and If So, How?: A Boundedly Rational Approach to Multi-armed Bandits. *Journal of Economic Theory*, *78*(1), 130–156. https://doi.org/10.1006/jeth.1997.2347

Schlag, K. H. (1999). Which one should I imitate? *Journal of Mathematical Economics*, *31*(4), 493–522. https://doi.org/10.1016/S0304-4068(97)00068-2

Schmidt, M. F. H., & Tomasello, M. (2012). Young Children Enforce Social Norms. *Current Directions in Psychological Science*, *21*(4), 232–236. https://doi.org/10.1177/0963721412448659

Schneider, F., Ledermann, T., Rist, S., & Fry, P. (2009). Social learning processes in Swiss soil protection - The "From Farmer - To Farmer" project. *Human Ecology*, *4*, 475–489. https://doi.org/10.1007/s10745-009-9262-1

Schoonenboom, J., & Johnson, R. B. (2017). How to Construct a Mixed Methods Research Design. *Kolner Zeitschrift Fur Soziologie Und Sozialpsychologie*, *69*, 107–131. https://doi.org/10.1007/s11577-017-0454-1

Schroeder, D. A., Graziano, W. G., Barclay, P., & Van Vugt, M. (2015). The Evolutionary Psychology of Human Prosociality. *The Oxford Handbook of Prosocial Behavior*, *October*, 1–29. https://doi.org/10.1093/oxfordhb/9780195399813.013.029

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The Constructive, Destructive, and Reconstructive Power of Social Norms. *Psychological Science*, *18*(5), 429–434. https://doi.org/doi: 10.1111/j.1467-9280.2007.01917.x.

Schweinfurth, M. K., & Taborsky, M. (2017). The transfer of alternative tasks in reciprocal cooperation. *Animal Behaviour, 131*, 35–41. https://doi.org/10.1016/j.anbehav.2017.07.007

Scotland, J. (2012). Exploring the philosophical underpinnings of research: Relating ontology and epistemology to the methodology and methods of the scientific, interpretive, and critical research paradigms. *English Language Teaching*, *5*(9), 9–16. https://doi.org/10.5539/elt.v5n9p9

Sharps, M. J., Villegas, A. B., Nunes, M. A., & Barber, T. L. (2002). Memory for animal tracks: A possible cognitive artifact of human evolution. *Journal of Psychology: Interdisciplinary and Applied*, *136*(5), 469–492. https://doi.org/10.1080/00223980209605545

Shrivastava, A., Jain, G., Kamble, S. S., & Belhadi, A. (2020). Sustainability through online renting clothing: Circular fashion fueled by instagram micro-celebrities. *Journal of Cleaner Production*, *278*, 123772. https://doi.org/10.1016/j.jclepro.2020.123772

Sigmund, K., De Silva, H., Traulsen, A., & Hauert, C. (2010). Social learning promotes institutions for governing the commons. *Nature*, *466*(7308), 861–863. https://doi.org/10.1038/nature09203

Silva, A. S., & Mace, R. (2014). Cooperation and conflict: Field experiments in Northern Ireland. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1792). https://doi.org/10.1098/rspb.2014.1435

Simon, H. A. (1990). A mechanism for social selection and successful altruism. *Science*. https://doi.org/10.1126/science.2270480

Simpson, B. J. K., & Radford, S. K. (2012). Consumer Perceptions of Sustainability: A Free Elicitation Study. *Journal of Nonprofit and Public Sector Marketing*, *24*(4), 272–291. https://doi.org/10.1080/10495142.2012.733654

Singh, M., Acerbi, A., Caldwell, C. A., Danchin, É., Isabel, G., Molleman, L., Scott-Phillips, T., Tamariz, M., van den Berg, P., van Leeuwen, E. J. C., & Derex, M. (2021). Beyond social learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1828). https://doi.org/10.1098/rstb.2020.0050

Skyrms, B. (2004). The Stag Hunt and the Evolution of Social Structure. In *Cambridge University Press*. Cambridge University Press. https://doi.org/10.1086/423724

Smaldino, P. E., Schank, J. C., & McElreath, R. (2013). Increased Costs of Cooperation Help Cooperators in the Long Run. *The American Naturalist*, *181*(4), 451–463. https://doi.org/10.1086/669615

Smith, B. (2018). Generalizability in qualitative research: misunderstandings, opportunities and recommendations for the sport and exercise sciences. *Qualitative Research in Sport, Exercise and Health*, *10*(1), 137–149. https://doi.org/10.1080/2159676X.2017.1393221

Smith, D. (2020). Cultural group selection and human cooperation: a conceptual and empirical review. *Evolutionary Human Sciences*, *2*, 1–29. https://doi.org/10.1017/ehs.2020.2

Smith, J. E. (2014). Hamilton's legacy: Kinship, cooperation and social tolerance inmammalian groups. *Animal Behaviour*, *92*, 291–304. https://doi.org/10.1016/j.anbehav.2014.02.029

Smith, J. E., Van Horn, R. C., Powning, K. S., Cole, A. R., Graham, K. E., Memenis, S. K., & Holekamp, K. E. (2010). Evolutionary forces favoring intragroup coalitions among spotted hyenas and other animals. *Behavioral Ecology*, *21*(2), 284–303. https://doi.org/10.1093/beheco/arp181

Smith, M. (2019). *Concern for the environment at record highs*. YouGov. https://yougov.co.uk/topics/politics/articles-reports/2019/06/05/concern-environment-record-highs

Smith, R. H., Powell, C. A. J., Combs, D. J. Y., & Schurtz, D. R. (2009). Exploring the When and Why of Schadenfreude. *Social and Personality Psychology Compass*, *3*(4), 530–546. https://doi.org/10.1111/j.1751-9004.2009.00181.x

Smith, R. P., Doiron, A., Muzquiz, R., Fortoul, M. C., Haas, M., Abraham, T., Quinn, R. J., Barraza, I., Chowdhury, K., & Nemzer, L. R. (2019). The public and private benefit of an impure public good determines the sensitivity of bacteria to population collapse in a snowdrift game. *Environmental Microbiology*, *21*(11), 4330–4342. https://doi.org/10.1111/1462-2920.14796

Smith, S., Windmeijer, F., & Wright, E. (2015). Peer effects in charitable giving: Evidence from the (Running) field. *Economic Journal*, *125*(585), 1053–1071. https://doi.org/10.1111/ecoj.12114

Smukalla, S., Caldara, M., Pochet, N., Beauvais, A., Guadagnini, S., Yan, C., Vinces, M. D., Jansen, A., Prevost, M. C., Latgé, J. P., Fink, G. R., Foster, K. R., & Verstrepen, K. J. (2008). FLO1 Is a Variable Green Beard Gene that Drives Biofilm-like Cooperation in Budding Yeast. *Cell*, *135*(4), 726–737. https://doi.org/10.1016/j.cell.2008.09.037

Sosis, R., & Bressler, E. R. (2007). Longevity : A Test of the Costly Signaling Theory of Religion. *Cross-Cultural Research*, *37*(2), 211–239. https://doi.org/10.1177/1069397103251426

Souza, M. O., Pacheco, J. M., & Santos, F. C. (2009). Evolution of cooperation under N-person snowdrift games. *Journal of Theoretical Biology*, *260*(4), 581–588. https://doi.org/10.1016/j.jtbi.2009.07.010

Srivastava, J., Espinoza, F., & Fedorikhin, A. (2009). Coupling and decoupling of unfairness and anger in ultimatum bargaining. *Journal of Behavioral Decision Making*. https://doi.org/10.1002/bdm.631

Stern, P. C. (2011). Contributions of Psychology to Limiting Climate Change. *American Psychologist*, *66*(4), 303–314. https://doi.org/10.1037/a0023235

Stigler, S. (2008). Fisher and the 5% Level. *Chance*, *21*(4), 12–12. https://doi.org/10.1080/09332480.2008.10722926

Stubbersfield, J. M., Tehrani, J., & Flynn, E. (2018). Faking the News: Intentional Guided Variation Reflects Cognitive Biases in Transmission Chains Without Recall. *Cultural Science Journal*, *10*(1), 54. https://doi.org/10.5334/csci.109

Stubbersfield, J. M., Tehrani, J. J., & Flynn, E. G. (2017). Chicken Tumours and a Fishy Revenge: Evidence for Emotional Content Bias in the Cumulative Recall of Urban Legends. *Journal of Cognition and Culture*, *17*(1–2), 12–26. https://doi.org/10.1163/15685373-12342189

Su, W., Potenza, M. N., Zhang, Z., Hu, X., Gao, L., & Wang, Y. (2018). Do individuals with problematic and non-problematic internet game use differ in cooperative behaviors with partners of different social distances? Evidence from the Prisoner's Dilemma and Chicken Game. *Computers in Human Behavior*, *87*(February), 363–370. https://doi.org/10.1016/j.chb.2018.05.040

Sullivan, G. M., & Feinn, R. (2012). Using Effect Size—or Why the P Value Is Not Enough. *Journal of Graduate Medical Education*, *4*(3), 279–282. https://doi.org/10.4300/jgme-d-12-00156.1

Szolnoki, A., & Chen, X. (2018). Competition and partnership between conformity and payoff-based imitations in social dilemmas. *New Journal of Physics*, *20*(9), 093008. https://doi.org/10.1088/1367-2630/aade3c

Tabi, A., Hille, S. L., & Wüstenhagen, R. (2014). What makes people seal the green power deal? - Customer segmentation based on choice experiment in Germany. *Ecological Economics*, *107*, 206–215. https://doi.org/10.1016/j.ecolecon.2014.09.004

Takahashi, T., & Ihara, Y. (2019). Cultural and evolutionary dynamics with best-of-k learning when payoffs are uncertain. *Theoretical Population Biology*, *128*, 27–38. https://doi.org/10.1016/j.tpb.2019.05.004

Tang, K. H. D. (2018). Correlation between sustainability education and engineering students' attitudes towards sustainability. *International Journal of Sustainability in Higher Education*, *19*(3), 459–472. https://doi.org/10.1108/IJSHE-08-2017-0139

Taylor, P. D., & Jonker, L. B. (1978). Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, *40*, 145–156. https://doi.org/10.1016/0025-5564(78)90077-9

Tehrani, J. J., Flynn, E. G., & Stubbersfield, J. M. (2015). Serial killers, spiders and cybersex: social and survival information bias in the transmission of urban legends. *British Journal of Psychology*, *106*(2), 288–307. https://doi.org/10.1063/1.2756072

TheGreenAge. (2019). *Is Renewable Energy Really More Expensive?* TheGreenage. https://www.thegreenage.co.uk/is-renewable-energy-really-more-expensive/

Theofanopoulou, C., Gastaldon, S., O'Rourke, T., Samuels, B. D., Messner, A., Martins, P. T., Delogu, F., Alamri, S., & Boeckx, C. (2017). Self-domestication in Homo sapiens: Insights from comparative genomics. *PLoS ONE*, *12*(10), 5–7. https://doi.org/10.1371/journal.pone.0185306

Thøgersen, J. (2008). Social norms and cooperation in real-life social dilemmas. *Journal of Economic Psychology*, *29*(4), 458–472. https://doi.org/10.1016/j.joep.2007.12.004

Thomas, M. G., Ji, T., Wu, J., He, Q., Tao, Y., & Mace, R. (2018). Kinship underlies costly cooperation in mosuo villages. *Royal Society Open Science*, *5*(2). https://doi.org/10.1098/rsos.171535

Thondhlana, G., & Hlatshwayo, T. N. (2018). Pro-environmental behaviour in student residences at Rhodes University, South Africa. *Sustainability (Switzerland)*, *10*(8), 1–19. https://doi.org/10.3390/su10082746

Tiefenbeck, V., Staake, T., Roth, K., & Sachs, O. (2013). For better or for worse? Empirical evidence of moral licensing in a behavioral energy conservation campaign. *Energy Policy*, *57*, 160–171. https://doi.org/10.1016/j.enpol.2013.01.021

Toelch, U., Bruce, M. J., Meeus, M. T. H., & Reader, S. M. (2010). Humans copy rapidly increasing choices in a multiarmed bandit problem. *Evolution and Human Behavior*, *31*(5), 326–333. https://doi.org/10.1016/j.evolhumbehav.2010.03.002

Toelch, U., Bruce, M. J., Newson, L., Richerson, P. J., & Reader, S. M. (2013). Individual consistency and flexibility in human social information use. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1776). https://doi.org/10.1098/rspb.2013.2864

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*. https://doi.org/10.1017/s0140525x0003123x

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two Key Steps in the Evolution of Human Cooperation. *Current Anthropology*. https://doi.org/10.1086/668207

Toyokawa, W., Whalen, A., & Laland, K. N. (2019). Social learning strategies regulate the wisdom and madness of interactive crowds. *Nature Human Behaviour*, *3*(2), 183–193. https://doi.org/10.1038/s41562-018-0518-x

Tripathi, A., & Singh, M. P. (2016). Determinants of sustainable/green consumption: A review. *International Journal of Environmental Technology and Management*, *19*(3–4), 316–358. https://doi.org/10.1504/IJETM.2016.082258

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology*, *46*(1), 35–57. https://doi.org/10.1086/406755

Truskanov, N., Emery, Y., & Bshary, R. (2020). Juvenile cleaner fish can socially learn the consequences of cheating. *Nature Communications*, *11*(1), 1–9. https://doi.org/10.1038/s41467-020-14712-3

Trut, L. (1999). Early Canid Domestication: The Farm-Fox Experiment. *American Scientist*. https://doi.org/10.1511/1999.20.813

Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in Psychology*, *8*(AUG). https://doi.org/10.3389/fpsyg.2017.01454

Vale, G. L., Flynn, E. G., Kendal, J., Rawlings, B., Hopper, L. M., Schapiro, S. J., Lambeth, S. P., & Kendal, R. L. (2017). Testing differential use of payoff-biased social learning strategies in children and chimpanzees. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1868), 1–9. https://doi.org/10.1098/rspb.2017.1751

van de Vyver, J., & Abrams, D. (2015). Testing the prosocial effectiveness of the prototypical moral emotions: Elevation increases benevolent behaviors and outrage increases justice behaviors. *Journal of Experimental Social Psychology*. https://doi.org/10.1016/j.jesp.2014.12.005

van de Vyver, J., & Leite, A. C. (2022). A longitudinal and experimental test of extended intergroup contact on social inclusion. *Unpublished Manuscript*.

van den Berg, P., Dewitte, P., Aertgeerts, I., & Wenseleers, T. (2020). How the incentive to contribute affects contributions in the one-shot public goods game. *Scientific Reports*, *10*(1), 8–12. https://doi.org/10.1038/s41598-020-75729-8

van den Berg, P., Molleman, L., Junikka, J., Puurtinen, M., & Weissing, F. J. (2015). Human cooperation in groups: Variation begets variation. *Scientific Reports*, *5*, 1–7. https://doi.org/10.1038/srep16144 van den Berg, P., Molleman, L., & Weissing, F. J. (2015). Focus on the success of others leads to selfish behavior. *Proceedings of the National Academy of Sciences*, *112*(9), 2912–2917. https://doi.org/10.1073/pnas.1417203112

Van Lange, P. A. M., Joireman, J., Parks, C. D., & Van Dijk, E. (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes*, *120*(2), 125–141. https://doi.org/10.1016/j.obhdp.2012.11.003

van Leeuwen, E. J. C., Acerbi, A., Kendal, R. L., Tennie, C., & Haun, D. B. M. (2016). A reappreciation of 'conformity.' *Animal Behaviour*, *122,* e5–e10. https://doi.org/10.1016/j.anbehav.2016.09.010

van Leeuwen, E. J. C., DeTroy, S. E., Kaufhold, S. P., Dubois, C., Schütte, S., Call, J., & Haun, D. B. M. (2021). Chimpanzees behave prosocially in a group-specific manner. *Science Advances*, *7*(9), 1–9. https://doi.org/10.1126/sciadv.abc7982

Verpooten, J., & Dewitte, S. (2017). The Conundrum of Modern Art: Prestige-Driven Coevolutionary Aesthetics Trumps Evolutionary Aesthetics among Art Experts. *Human Nature*, *28*(1), 16–38. https://doi.org/10.1007/s12110-016-9274-7

Vincent, A. (2017). Punishment and status in collective action: How status hierarchies foster optimal punishment use. *Sociology Compass*, *11*(6), 1–11. https://doi.org/10.1111/soc4.12478

Vine, D., Buys, L., & Morris, P. (2013). The Effectiveness of Energy Feedback for Conservation and Peak Demand: A Literature Review. *Open Journal of Energy Efficiency*, *2*, 7–15. https://doi.org/10.4236/ojee.2013.21002

Volk, S., Thöni, C., & Ruigrok, W. (2011). Personality, personal values and cooperation preferences in public goods games: A longitudinal study. *Personality and Individual Differences*, *50*(6), 810–815. https://doi.org/10.1016/j.paid.2011.01.001

von Korff, Y., Daniell, K. A., Moellenkamp, S., Bots, P., & Bijlsma, R. M. (2012). Implementing participatory water management: Recent advances in theory, practice, and evaluation. *Ecology and Society*, *17*(1), 30. https://doi.org/10.5751/ES-04733-170130

von Rueden, C., Gurven, M., Kaplan, H., & Stieglitz, J. (2014). Leadership in an Egalitarian Society. *Human Nature*. https://doi.org/10.1007/s12110-014-9213-4

Walton, T., & Austin, D. M. (2011). Pro-environmental behavior in an urban social structural context. *Sociological Spectrum*, *31*(3), 260–281. https://doi.org/10.1080/02732173.2011.557037

Waring, T., & Acheson, J. (2018). Evidence of cultural group selection in territorial lobstering in Maine. *Sustainability Science*, *13*(1), 21–34. https://doi.org/10.1007/s11625-017-0501-x

Waring, T. M., Goff, S. H., & Smaldino, P. E. (2017). The coevolution of economic institutions and sustainable consumption via cultural group selection. *Ecological Economics*, *131*, 524–532. https://doi.org/10.1016/j.ecolecon.2016.09.022

Watson, R., Morgan, T. J. H., Kendal, R. L., Vyver, J. van de, & Kendal, J. (2021). Social Learning Strategies and Cooperative Behaviour : Evidence of Payoff Bias , but Not Prestige or Conformity , in a Social Dilemma Game. *Games*, *12*(4), 89. https://doi.org/10.3390/g12040089

Watts, L. L., Ness, A. M., Steele, L. M., & Mumford, M. D. (2018). Learning from stories of leadership: How reading about personalized and socialized politicians impacts performance on an ethical decision-making simulation. *Leadership Quarterly*, *29*(2), 276–294. https://doi.org/10.1016/j.leaqua.2017.04.004

Webster, M. M., & Laland, K. N. (2008). Social learning strategies and predation risk: Minnows copy only when using private information would be costly. *Proceedings of the Royal*

Society B: Biological Sciences, *275*(1653), 2869–2876. https://doi.org/10.1098/rspb.2008.0817

Webster, M. M., & Laland, K. N. (2011). Reproductive state affects reliance on public information in sticklebacks. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1705), 619–627. https://doi.org/10.1098/rspb.2010.1562

Wechsler, H., Nelson, T. E., Lee, J. E., Seibring, M., Lewis, C., & Keeling, R. P. (2003). Perception and reality: a national evaluation of social norms marketing interventions to reduce college students' heavy alcohol use. *Journal of Studies on Alcohol, 64*(4), 484–494. https://doi.org/10.15288/jsa.2003.64.484

Weinstein, Y., Bugg, J. M., & Roediger, H. L. (2008). Can the survival recall advantage be explained by basic memory processes? *Memory and Cognition*, *36*(5), 913–919. https://doi.org/10.3758/MC.36.5.913

West, S. A., & Gardner, A. (2010). Altruism, spite, and greenbeards. In *Science*. https://doi.org/10.1126/science.1178332

West, S. A., Griffin, A. S., & Gardner, A. (2007a). Evolutionary Explanations for Cooperation. *Current Biology*, *17*(16), 661–672. https://doi.org/10.1016/j.cub.2007.06.004

West, S. A., Griffin, A. S., & Gardner, A. (2007b). Social semantics: Altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology*, *20*(2), 415–432. https://doi.org/10.1111/j.1420-9101.2006.01258.x

White, K., Habib, R., & Hardisty, D. J. (2019). How to SHIFT consumer behaviors to be more sustainable: A literature review and guiding framework. *Journal of Marketing*, *83*(3), 22–49. https://doi.org/10.1177/0022242919825649

Whitehead, H., & Richerson, P. J. (2009). The evolution of conformist social learning can cause population collapse in realistically variable environments. *Evolution and Human Behavior, 30*(4), 261–273. https://doi.org/10.1016/j.evolhumbehav.2009.02.003

Whiten, A. (2019a). Conformity and over-imitation: An integrative review of variant forms of hyper-reliance on social learning. *Advances in the Study of Behavior, 51*, 31–75. https://doi.org/10.1016/bs.asb.2018.12.003

Whiten, A. (2019b). Cultural Evolution in Animals. *Annual Review of Ecology, Evolution, and Systematics, 50,* 27–48. https://doi.org/10.1146/annurev-ecolsys-110218-025040

Williamson, R. A., & Meltzoff, A. N. (2011). Own and others' prior experiences influence children's imitation of causal acts. *Cognitive Development*, *26*(3), 260–268. https://doi.org/10.1016/j.cogdev.2011.04.002

Wilson, E. O. (1975). *Sociobiology: The New Synthesis*. Cambridge, Mass. : Belknap Press of Harvard University Press. https://doi.org/10.2307/589826

Wit, A. P., & Wilke, H. A. M. (1992). The effect of social categorization on cooperation in three types of social dilemmas. *Journal of Economic Psychology*, *13*(1), 135–151. https://doi.org/10.1016/0167-4870(92)90056-D

Wood, L. A., Kendal, R. L., & Flynn, E. G. (2013). Whom do children copy? Model-based biases in social learning. *Developmental Review*, *33*(4), 341–356. https://doi.org/10.1016/j.dr.2013.08.002

Wood, L. A., Kendal, R. L., & Flynn, E. G. (2015). Does a peer model's task proficiency influence children's solution choice and innovation? *Journal of Experimental Child Psychology*, *139*, 190–202. https://doi.org/10.1016/j.jecp.2015.06.003

Wood, N. T., & Herbst, K. C. (2007). Political star power and political parties: Does celebrity endorsement win first-time votes? *Journal of Political Marketing*, *6*(2–3), 141–158. https://doi.org/10.1300/J199v06n02_08

Wood, R. I., Kim, J. Y., & Li, G. R. (2016). Cooperation in rats playing the iterated Prisoner's Dilemma game. *Animal Behaviour*, *114*, 27–35. https://doi.org/10.1016/j.anbehav.2016.01.010

Wu, J., Balliet, D., & van Lange, P. A. M. (2016). Reputation, Gossip, and Human Cooperation. *Social and Personality Psychology Compass*, *10*(6), 350–364. https://doi.org/10.1111/spc3.12255

WWF. (2021). *Leonardo DiCaprio. Founder, Leonardo DiCaprio Foundation*. Leadership. https://doi.org/10.9774/gleaf.978-1-909493-73-5_276

Wynne-Edwards, V. C. (1963). Intergroup selection in the evolution of social systems. *Nature*. https://doi.org/10.1038/200623a0

Xu, J., Dowman, M., & Griffiths, T. L. (2013). Cultural transmission results in convergence towards colour term universals. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1758). https://doi.org/10.1098/rspb.2012.3073

Xu, L., Ling, M., & Wu, Y. (2018). Economic incentive and social influence to overcome household waste separation dilemma: A field intervention study. *Waste Management*, *77*, 522–531. https://doi.org/10.1016/j.wasman.2018.04.048

Xygalatas, D., Mitkidis, P., Fischer, R., Reddish, P., Skewes, J., Geertz, A. W., Roepstorff, A., & Bulbulia, J. (2013). Extreme Rituals Promote Prosociality. *Psychological Science*, *24*(8), 1602–1605. https://doi.org/10.1177/0956797612472910

Yamin, P., Fei, M., Lahlou, S., & Levy, S. (2019). Using social norms to change behavior and increase sustainability in the real world: A systematic review of the literature. *Sustainability*, *11*(20), 5847. https://doi.org/10.3390/su11205847

Yılmaz, N., & Erkal, S. (2017). Determining Undergraduate Students' Environmental Attitude. *European Journal of Sustainable Development*, *6*(4), 137–146. https://doi.org/10.14207/ejsd.2017.v6n4p137

Yoeli, E., Hoffman, M., Rand, D. G., & Nowak, M. A. (2013). Powering up with indirect reciprocity in a large-scale field experiment. *Proceedings of the National Academy of Sciences*, *110*(Supplement_2), 10424–10429. https://doi.org/10.1073/pnas.1301210110

Young, W., Hwang, K., McDonald, S., & Oates, C. J. (2010). Sustainable consumption: Green consumer behaviour when purchasing products. *Sustainable Development*, *18*(1), 20–31. https://doi.org/10.1002/sd.394

Zack, E. S., Kennedy, J. M., & Long, J. S. (2019). Can nonprobability samples be used for social science research? A cautionary tale. *Survey Research Methods*, *13*(2), 215–227. https://doi.org/10.18148/srm/2019.v13i2.7262

Zanella, M., Vitriolo, A., Andirko, A., Martins, P. T., Sturm, S., O'Rourke, T., Laugsch, M., Malerba, N., Skaros, A., Trattaro, S., Germain, P. L., Mihailovic, M., Merla, G., Rada-Iglesias, A., Boeckx, C., & Testa, G. (2019). Dosage analysis of the 7q11.23 Williams region identifies BAZ1B as a major human gene patterning the modern human face and underlying self-domestication. *Science Advances*, *5*(12), eaaw7908. https://doi.org/10.1126/sciadv.aaw7908

Zefferman, M. R. (2018). Cultural multilevel selection suggests neither large or small cooperative agreements are likely to solve climate change without changing the game. *Sustainability Science*, *13*(1), 109–118. https://doi.org/10.1007/s11625-017-0488-3

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics*, *6*(3), 299–310. https://doi.org/10.1023/A:1026277420119

Zhang, L., & Ortmann, A. (2016). Pro-social or anti-social, or both? A within- and between-subjects study of social preferences. *Journal of Behavioral and Experimental Economics*, *62*, 23–32. https://doi.org/10.1016/j.socec.2016.03.001

Zhong, L. X., Qiu, T., & Xu, J. R. (2008). Heterogeneity improves cooperation in continuous snowdrift game. *Chinese Physics Letters*, *25*(6), 2315–2318. https://doi.org/10.1088/0256-307X/25/6/107

Zimmermann, M. G., & Eguíluz, V. M. (2005). Cooperation, social networks, and the emergence of leadership in a prisoner's dilemma with adaptive local interactions. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *72*(5). https://doi.org/10.1103/PhysRevE.72.056118

Zizzo, D. J., & Fleming, P. (2011). Can experimental measures of sensitivity to social pressure predict public good contribution? *Economics Letters*, *111*(3), 239–242. https://doi.org/10.1016/j.econlet.2011.02.021

Zizzo, D. J., & Oswald, A. J. (2001). Are People Willing to Pay to Reduce Others' Incomes? *Annales d'Économie et de Statistique*, *63/64*, 39–65. https://doi.org/10.2307/20076295

Appendices

Appendix 3.1 – Power analysis for study 1

Introduction

When evaluating the required sample size for quantitative research, one approach is to use an apriori power analysis. Statistical power refers to the probability of detecting an effect if a true effect exists (J. Cohen, 1992). That is, with sufficient statistical power, one can be confident that the results obtained from statistical analysis have not arisen spuriously or are highly uncertain. The realised power of an experiment is dependent both on the strength of the effect being tested and on the sample size. Given that it is not possible to modify the strength of effects, the simplest way to increase statistical power is to increase the sample size. In his original paper, (J. Cohen, 1992) suggested that the goal for experimental research should be to achieve a statistical power of 0.8. This means that, as a researcher, you have an 80% of detecting a significant effect, assuming one is present. He provided a table listing the sample size requirements for many common statistical effects to achieve this power under small, medium and large effects.

This approach works generally well for cases where researchers use traditional null hypothesis testing to evaluate their results. Under this framework, the null hypothesis (no effect) is rejected if the data observed would be highly unlikely (conventionally, less than 5% likely (Stigler, 2008)) if the null hypothesis were true. Sufficient statistical power is very important to ensure that one does not reject the null hypothesis when there in reality is no effect present (type 1 error) or fail to reject the null hypothesis when there is an effect present (type 2 error).

Statistical power does not have a clear application for Bayesian statistics, which do not use P values, nor do they seek to reject any null hypothesis. Instead, the goal of a Bayesian analysis is to use both prior information and the data put into the model to build posterior distributions which represent a series of plausible values which could have resulted in the data you collected. The key difference between this and null hypothesis testing, is that no one point of the posterior distribution is of particular interest. Instead, the entire distribution, and the uncertainty it represents, is of equal importance for answering the research question (for a further discussion on Bayesian approaches, see (McElreath, 2016)). Therefore, because the answers gleaned from Bayesian analyses are not binary (reject or accept) then it is not useful to consider statistical power in the manner described above.

Nevertheless, in a practical sense, it is still important to have some idea about how large the sample needs to be to accurately model the effects being investigated. Although Bayesian analyses do not risk type 1 or type 2 errors, a small sample will produce posterior estimates of high uncertainty meaning inferring the strength of expected pattern of results can be problematic. That is, if the effect in the data, or the amount of data, is not sufficiently large, then the model cannot become confident of the parameter associated with this effect. In this case, a larger sample would be required to increase this certainty.

In the proceeding sections, the question of sample size is investigated using simulated data which precisely resemble the type and format of the data which will be collected in the experiment. Experimental runs are simulated with each of the three social learning strategies of interest, conformity, prestige and payoff, and Bayesian regressions are then fitted to this data. The goal is to investigate how the increase in sample size affects the uncertainty of the social learning parameter which is known to be present.

The model

Description of the model

This model seeks to emulate a typical public goods game commonly seen in experimental scenarios and other simulation work (Doebeli & Hauert, 2005b; Neugebauer et al., 2009), which are usually based on the prisoner’s dilemma (Rapoport, 1974). In these games, across multiple rounds, agents choose how much of their personal endowment to donate to a public fund, which is then multiplied by some factor and distributed evenly amongst all players, irrespective of how much they contributed. The game functions such that those who do not cooperate achieve greater payoffs than those that do, despite the greatest net payoff only being achievable if everyone cooperates. As a result, and following predictions from game theory, typical findings in both experimental and theoretical work is that cooperation declines rapidly as the games progress.

In this model, groups of 4 agents are placed into a group and play 6 rounds of a public goods game. 6 rounds are used, because this is how many rounds are played in the follow up experiment. One of these agents is randomly assigned to be “prestigious”. In each round, agents receive an initial 10 monetary units and may decide to donate any amount of this to the public good. Once all agents have donated, the pot is multiplied by 2 and redistributed between the agents. Their payoff for that round is the sum of how much they kept for themselves plus what they earned from the round. For example, if a participant donated 3 and kept 7 and received 12 from the pot, their payoff would be 21 for that round.

Additionally, this model also considers the impact of social learning strategies. Following each round, agents are granted the opportunity to modify their behaviour. This can happen in one of three ways. Through conformity, the agent compares their contribution to the mean contribution of the group and if they donated more or less than this value, their next contribution moves 1 point towards this value. For example, if an agent donated 3 units but the mean was 5, then their next contribution would be 4. For payoff-based learning, the agent looks at the agent with the highest cumulative payoff and exactly copies their previous contribution. Similarly, agents may exactly copy the prestigious individual irrespective of their payoff which can result in copying both higher and lower cooperation values. Based upon the findings of other theoretical models that prestigious individuals are more cooperative than non-prestigious individuals (Henrich et al., 2015), there is also an additional probability that the prestigious individual might increase their contribution by 1 after the copying has taken place (see below). At each round, it is only possible for an agent to use one of these strategies, but they are not fixed in what strategy they use.

For this model, it is assumed that these are the only ways that agents may modify their behaviour. The copying employed in this model is assumed to be perfect and the only mechanism available for behaviour to change. This ignores other potentially important mechanisms such as reciprocity (Trivers, 1971), asocial learning or other random mutations across rounds which could plausibly affect participants cooperation in experimental contexts. Simplifying the model in this way means it is sufficient to estimate a minimum sample size, but it cannot faithfully represent the behavioural noise seen in experimental games.

Model parameters

Each agent's starting cooperation is determined by sampling from a normal distribution with a mean of 5 and a standard deviation of 1.5 (figure 3.4). To prevent the sampling of values with decimal places, the generated values are rounded to the nearest whole number. This prior distribution of cooperation is motivated by findings typically observed in experimental public goods games. On average, participants normally begin donating half of what they could (5 in this case) but between participant variation is generally very high. Therefore, this starting distribution captures the general pattern observed in experimental games (Neugebauer et al., 2009).

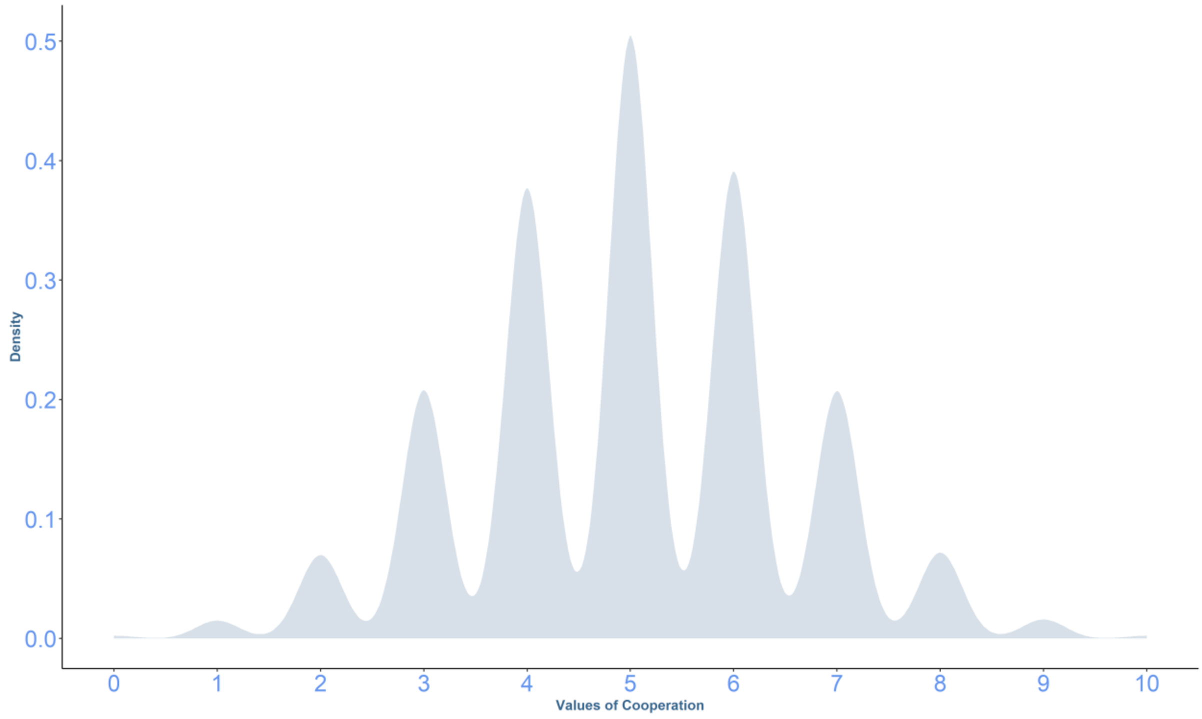


Figure 3.4. *Starting distribution of cooperation sampled from by all agents. Normal ~ (5, 1.5). Values are rounded to the nearest whole number.*

For the purposes of this model, the prestigious individual is assumed to be slightly more cooperative than the rest of the group. This is in light of theoretical and ethnographic work which suggest that an individual with prestige may be selected, or otherwise inclined, to behave more cooperatively (Henrich et al., 2015; Lenfesty & Morgan, 2019; von Rueden et al., 2014). Therefore, at the end of each round, the prestigious agent can, with a certain probability, increase their cooperation by 1. This probability is shown in the equation below.

$$Modify \sim Binomial(1, p)$$

$$p \sim Normal\ (0.5\ , 0.1)$$

Modification is sampled from a binomial distribution with a normal distribution of probabilities with a mean 0.5 and standard deviation 0.1. In each case, a probability is taken from this distribution and a single binomial trial generated from the sampled probability. If the resultant value is 1, then the prestigious individual will increase their cooperation by 1.

As described above, agents may modify their behaviour according to one social learning strategy each round. The strategy employed is assumed to vary across rounds but the probability that each strategy is used is constant across all agents and the strategy chosen in the previous round has no influence on the choice of strategy for the following round. To do this, agents take 1 random sample from an 100 digit long multinomial distribution containing numbers from 1 to 4 (see figure 3.5). These numbers correspond to the three social learning strategies described above (conformity, payoff, and prestige) as well as an asocial option, where the agent will not modify their behaviour at all that round. It is set up this way, to both ensure that there are no order effects with respect to which strategies are sampled for first and the exact probability that these strategies are employed can be controlled to a high degree of accuracy. For example, if an agent always checked whether they should use conformity before payoff, this may produce a different outcome than if this order was reversed.

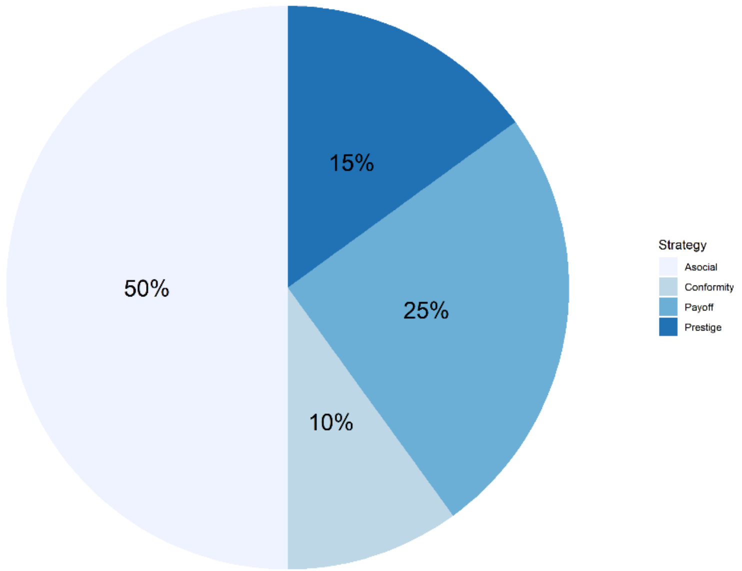


Figure 3.5. *Pie chart displaying the percentage of the distribution which is occupied by the 4 behaviour modification strategies. There is a 10% chance of using conformity, a 15% chance of using prestige, a 25% chance of using payoff and 50% chance of not changing behaviour (asocial).*

The percentage distribution of each strategy is motivated by their observed effect strengths in previous research. Payoff based social learning has been shown to be a strong influence in experimental cooperative contexts (Burton-Chellew, el Mouden, et al., 2017; Burton-Chellew et al., 2015) as participants will readily use payoff-based copying and reduce their contributions compared with frequency information. As such, the probability of modifying by payoff is relatively high compared to the other strategies. Conformity, on the other hand, has been shown to have much less influence on participants behaviour and so was assigned a far lower probability (Burton-Chellew, el Mouden, et al., 2017). Prestige has only seldom been investigated and although its effects are positive on cooperation in theoretical models (Henrich et al., 2015) and those of high status are copied in some cases (Kumru & Vesterlund, 2010) prestige remains an uncertain effect. As such, it was assigned a probability between conformity and payoff.

However, because the overall impact and use of social learning in experimental games is still relatively unknown, especially in relation to other strategies (but see (Bardsley & Sausgruber, 2005; Romano & Balliet, 2017), the chance of using any social learning was assumed to be only 50%. Though these probabilities have been selected somewhat arbitrarily, they permit a degree of uncertainty with regards to their effect on cooperation which is likely to remain true in the experimental context.

Analytical approach

The principal goal of this model is to evaluate how large the sample size needs to be in the experiment to give the greatest chance of reliably identifying the parameter values for the social learning strategies. This was because the project had limited funds to pay participants, so I did not want to use more participants than were required. To achieve this, the simulation model was repeated across a range of sample sizes and then compared to a run which uses an unachievably large sample size of 10000 participants (the "population" run). A sample of this size was sufficiently large to average out any effects of noise and highlight both the overall trend of cooperation given the probabilities described above but also permit extremely accurate parameter estimates from the Bayesian model. For each sample size, the following Bayesian model was fit to the data and the social learning parameters from the smaller sample runs were compared with those from the population run.

$$Cooperation_i \sim Normal\ (\mu\ , \sigma)$$

$$\mu_i = \ \alpha + \ \beta_{conformity} * Average\ donation + \ \beta_{prestige}$$

$$* \ Prestigious\ individual's\ cooperation + \ \beta_{payoff}$$

$$* Top\ scoring\ individual's\ cooperation\ + \ \beta_{prestigious\ individual}$$

$$* Prestigous\ individual$$

$$\alpha \sim Normal(5, 1.5)$$

$$\beta \sim Log - Normal(0, 1)$$

$$\sigma \sim Exponential(1)$$

Cooperation is described by its mean and variance and the mean is predicted by a linear model using four slope parameters. The parameters, $ß_{conformity}$, $ß_{prestige}$ $ß_{payoff}$ correspond to the three social learning strategies where they are multiplied by the social information from the previous round. This means that the cooperation at round *N* is predicted from the social information in round *N*-1. $ß_{prestigious\_individual}$ corresponds to being the prestigious agent which is modelled to be more cooperative. The prior for α makes use of the known starting distribution of cooperation while all ß parameters use a log-normal distribution to constrain the value to be positive. This is because the simulation only ever causes agents to move towards the observed cooperation and for the prestigious agent to be more cooperative, meaning the parameters will always be positive. Similarly, σ is given an exponential prior to constrain it to be positive. Once the models have been fit, samples will be drawn from the posterior distribution and the mean parameter estimates plus their prediction intervals will be displayed for each social learning parameters. Inevitably, the degree of uncertainty will be higher for smaller sample sizes, but if the parameters are very close to the population run, then it suggests that the sample size probably is sufficient to detect such an effect in the experiment. All analysis is conducted in R studio version 4.0.2, models fit using the *rethinking* package (which interfaces with *Rstan* and uses *MCMC* for sampling) (McElreath, 2020) and plots produced using *ggplot2*.

Results

In the proceeding section, I will detail the general observed trend in cooperation in the population run (*N* = 10000) before then detailing the results of the Bayesian models and how the parameter estimates differed across the sample sizes.

Effect on Cooperation

Following predictions from game theory and simulation work, it would be expected that cooperation will decline from its initial starting point (Doebeli & Hauert, 2005b). Though, this model runs for fewer generations than is typical of prisoner's dilemma games and has relatively relaxed payoff-based incentives to reduce cooperation (so the trend may not be as pronounced in other work), cooperation still declines from its starting point (figure 3.6).

The distribution at round 1 is expected, given that cooperation was sampled from a distribution with a tight peak around 5. A very small proportion of groups may have sampled exclusively low values while an equal proportion sampled exclusively high values. However, the majority will sample a mixture, resulting in the most common value being 5. The distribution at round 6 shows a notable shift from the starting point, but in both directions. The tails of the distribution have become longer at the upper and the lower ends. There are also noticeable peaks towards the centre of the distribution. Overall, cooperation does appear to have reduced, but the effect is not especially strong. This distribution at round 6 suggests several things. Firstly, because payoff learning was the most common way (besides doing nothing) that agents would modify their behaviour and payoff learning exclusively results in copying lower cooperation, this will likely cause cooperation to trend downwards. The weakness of the trend is likely due to prestige-based copying, which offers a way for cooperation to increase. The peaks in the distribution are also likely a result of conformity, which will pull groupmates in line with one another.

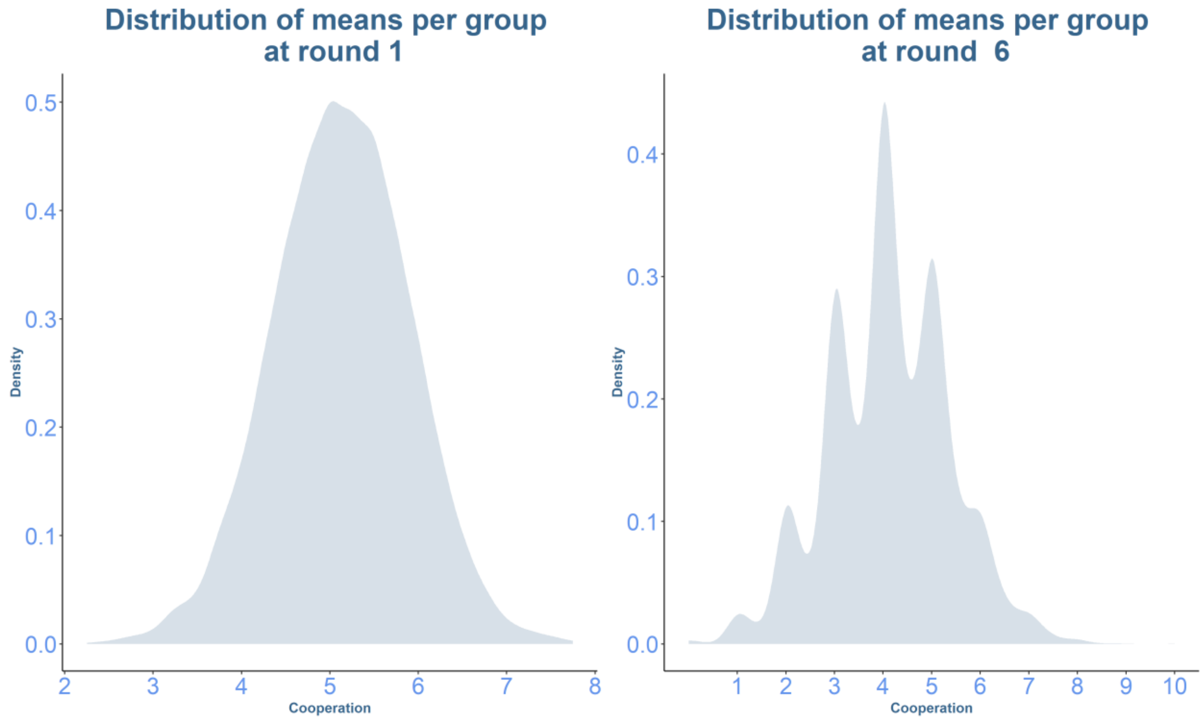


Figure 3.6. *Mean cooperation across all groups at round 1 (left) and round 6 (right).*

This may be clarified by figure 3.7, which shows a histogram displaying the frequency of each cooperation amount across the whole population. The overall pattern is the same, while some individuals maintained a high level of cooperation at the end of the game, the general distribution has shifted downwards. If this simulation were to run for more rounds, it is likely that cooperation would drop to 0.

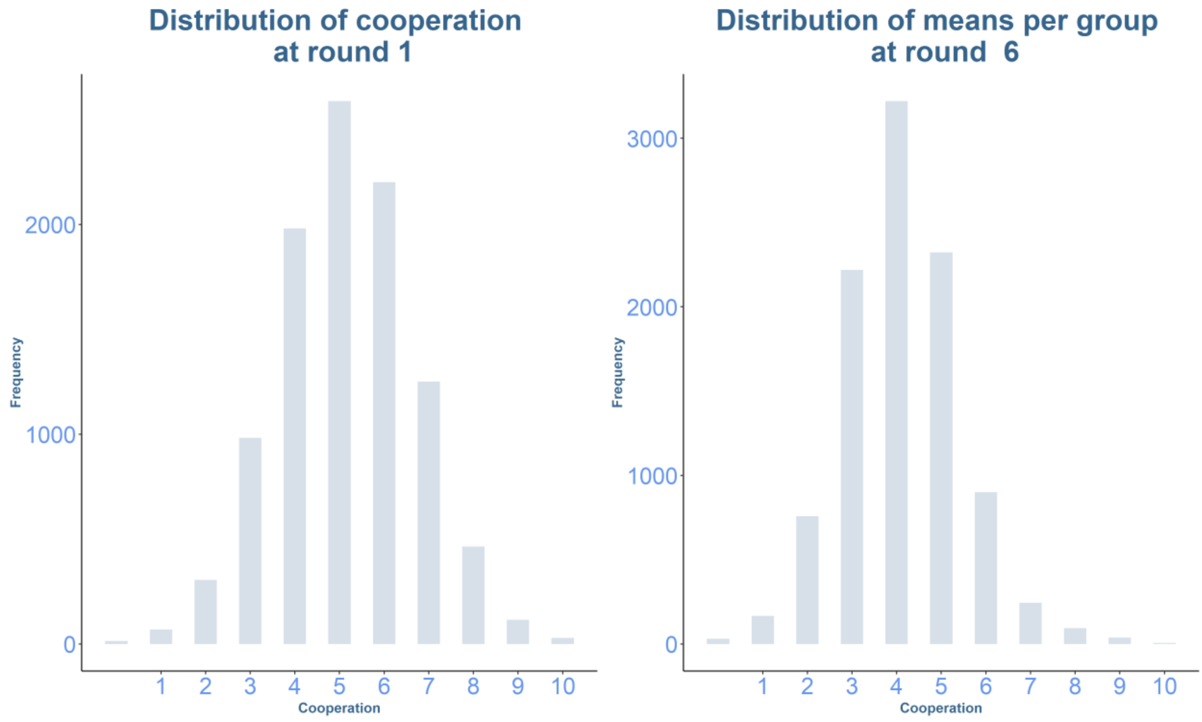


Figure 3.7. *Individual distribution of cooperation at round 1 compared to round 6.*

Fitting the models

A total of 5 Bayesian models were then fit to 5 runs of the simulation model using sample sizes of 5, 10, 15 and 20 groups of agents. Once completed, 12000 random samples of each of the four ß parameters were extracted and plotted against the population run to see how close they were. Though the uncertainty will decrease following increases in sample size, the primary interest is how close the posterior means are to one another and what proportion of the posterior distribution falls below 0.

It should be noted that, regardless of the strategy, the slopes should always be positive, if they are identified correctly. This may seem counterintuitive, given the trend in cooperation across the simulation was for it to reduce. The reason the slopes are positive is that, in every case, agents simply copy (or move towards) the cooperation that they socially observe, but never move away from the cooperation they observe. For example, if an agent observed the prestigious agent donating 10 and modified their behaviour through prestige learning, they would only ever increase their cooperation. This means, that if they observe high cooperation, then their cooperation will always increase but if they observe low cooperation, then it will always decrease. Therefore, for every strategy, lower values of observed cooperation will always predict lower values of agent cooperation while higher observed cooperation will always predict high participant cooperation. Hence, the beta slope will always be positive, which is why the parameters were assigned a log-normal prior. Although the models used log-normal priors for the beta parameters, this was relaxed to *Normal ~ (0,3)* for the smallest sample size, as the model had difficulty identifying parameters in this model. This difficulty in sampling can probably be explained by the patterns outlined below.

Conformity

Figure 3.8 shows the extracted estimates for conformity where the true value is 0.6. The mean of every slope parameter is positive, though only 15, 20 and 25 are reliably above 0. Therefore, the model was reasonably successful at predicting that an increase in the average donation resulted in an increase in agent cooperation and a sample size of 15 may be sufficient to achieve this. The parameter estimates (besides 30) were also very close to the true value within a reasonable degree of certainty. The sample run at 30 produces a strange result, in that it overlapped 0 and the mean was considerably less than the other estimates. This is likely explainable by sampling variation as conformity may have both been used less in those simulation runs, or the agents' cooperation may have started closer to the mean value (5) and moved less due to conformity. In any case, increasing the sample size has decreased the uncertainty associated with the parameter values and the posterior distribution always overlaps to true value. This is especially encouraging for conformity as it is the strategy that had the lowest probability of being employed and would be expected to be the hardest to detect.

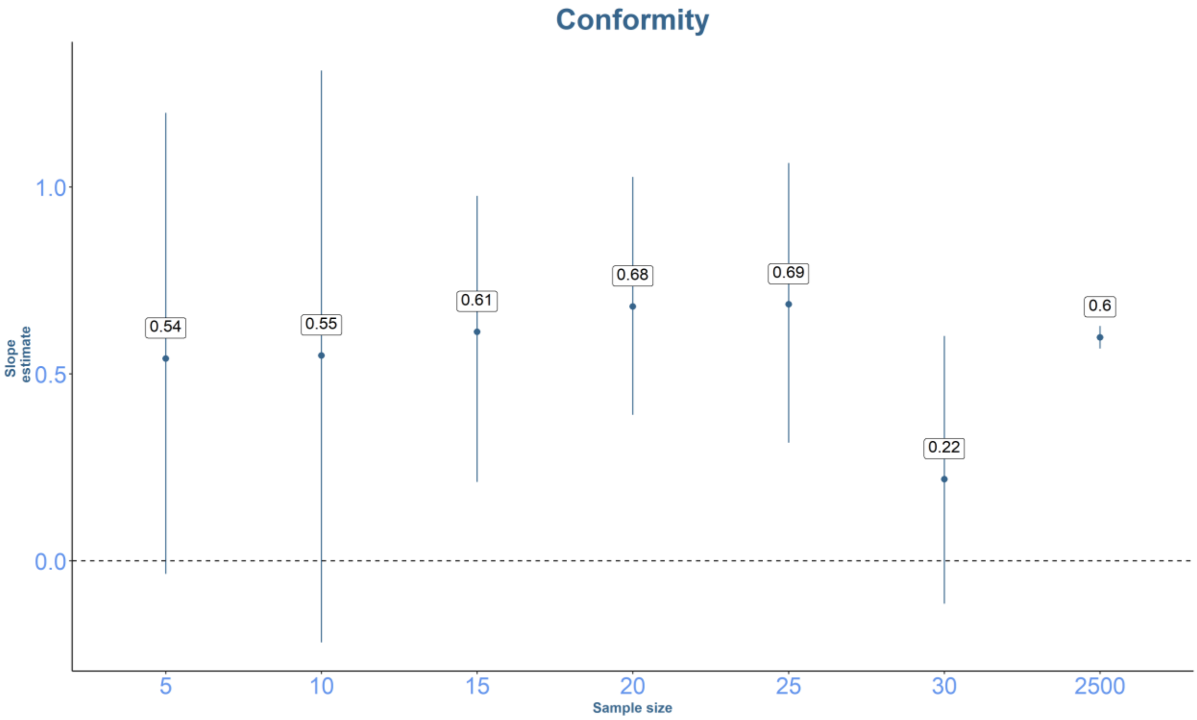


Figure 3.8. *Estimates for the slope parameter associated with conformist social learning. Points show the mean value of the posterior distribution (labelled above) and the lines represent the minimum and maximum values of the whole distribution.*

Prestige

The pattern for prestige is different to conformity (figure 3.9). The true value in this case was much smaller (0.14) but the population model become extremely certain in this estimation. From the remaining estimates, it is clear the model had difficulty detecting prestige learning reliably. The parameter estimates below 15 were extremely wide and overlapped 0 considerably. Past 15, the uncertainty decreased, but the estimates remained far away from the true value. Especially concerning is the estimate at 20 where the mean was negative. The only sample run which was reliably above 0 was 30 which itself, overestimated the parameter for prestige. Nevertheless, in some of the runs, the estimate was close to the true value and the degree of uncertainty was also similar past a sample size of 10.

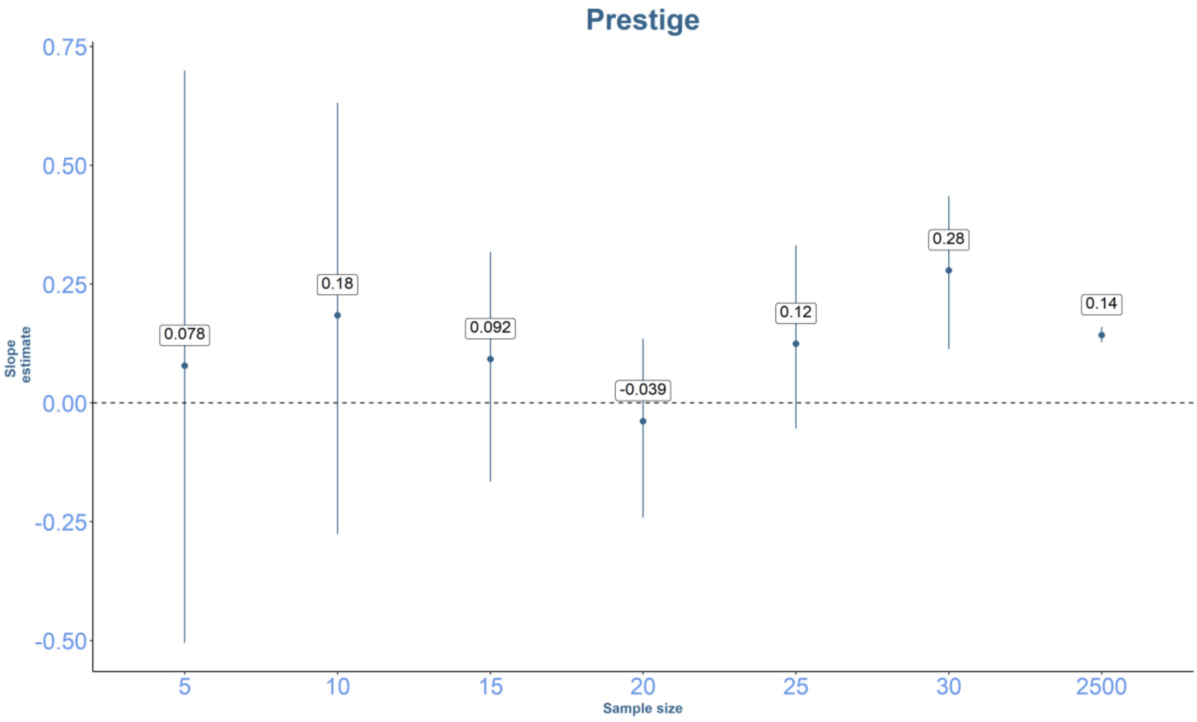


Figure 3.9. *Estimates for the slope parameter associated with prestige based social learning. Points show the mean value of the posterior distribution (labelled above) and the lines represent the minimum and maximum values of the whole distribution.*

Payoff

Compared to prestige and conformity, payoff was detected fairly reliably (figure 3.10). All of the parameter estimates (with exception of 5) were fairly close to the true value of 0.26 and the means were all positive with the majority of the posterior distribution being above 0. Increasing the sample size beyond 5 rapidly decreased the parameter uncertainty and all of the considered sample sizes could proxy the true parameter value reasonably accurately. None of the posterior means were negative meaning the predicted direction of the effects were correct.

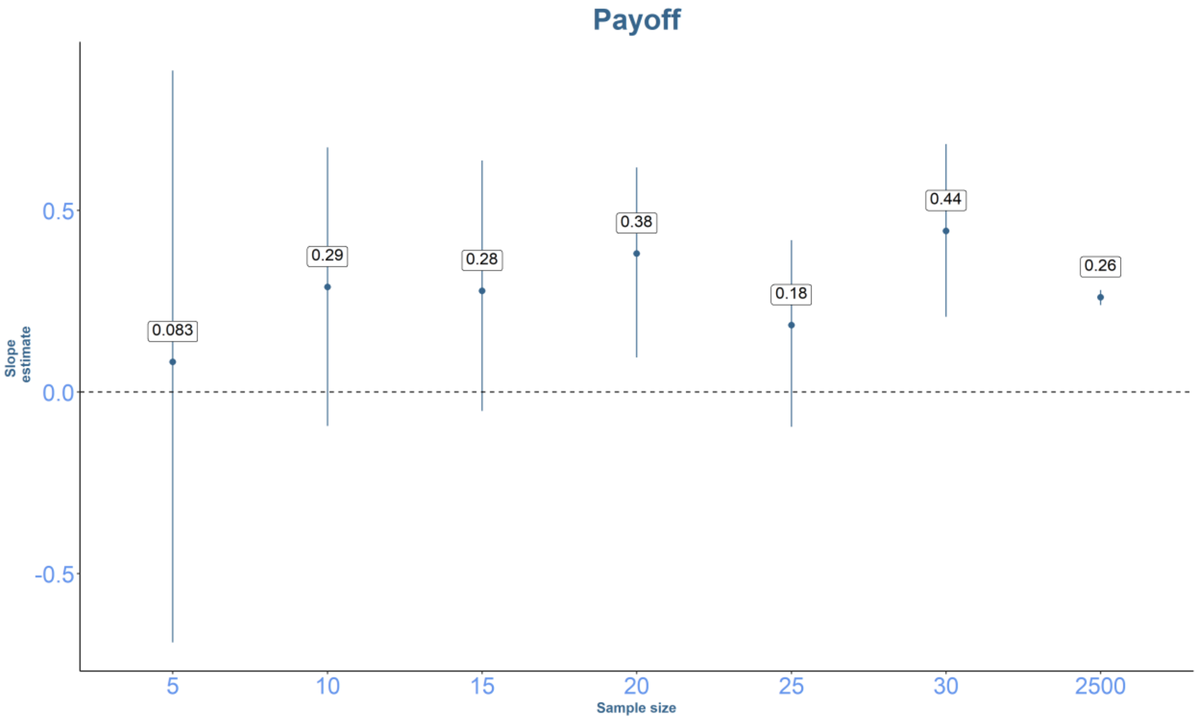


Figure 3.10. *Estimates for the slope parameter associated with payoff based social learning. Points show the mean value of the posterior distribution (labelled above) and the lines represent the minimum and maximum values of the whole distribution.*

Summary

Overall, a sample of size of around 20 seemed to be enough to successfully retrieve a parameter estimate close to the true value for conformity and payoff. Prestige was more difficult to detect, being the only strategy to have a negative posterior mean erroneously predicted and had the most estimates overlapping 0. However, even prestige showed a good decrease in parameter uncertainty with a sample of 15 and bar the estimate produced by 1 run, each estimate was in the correct direction and was close to the true value.

Another notable point is that the parameter values associated with prestige and payoff (0.14, 0.26) were considerably lower than that of conformity (0.60). This is a little curious, given that conformity was the least likely strategy to be employed. One explanation might be because of the way that the strategies function. Because the prestigious participant or the participant with the highest payoff cannot copy themselves, necessarily, in each round a maximum of only 3 participants can use payoff / prestige learning. For conformist learning, all 4 participants could theoretically modify their behaviour through conformity. For instance, if the donations in a round were 1,2, 9 and 10 (mean = 5.5) then every participant could change their donation in the next round. The result of this, is that the model would predict a relatively stronger effect for conformity than for payoff or prestige. Furthermore, with prestige, even if the participant does employ prestige-based learning, it is assumed that they only increase their cooperation by 1. Whereas for payoff learning, they directly copy the behaviour of the target individual. This might further explain why the parameter value associated with payoff learning is higher than that of prestige.

Improving the estimate for Prestige

Because prestige stood out as the most difficult strategy to reliably identify, a further statistical strategy was employed to try and improve this estimate. As described above, the issue with prestige is that one agent in every group will never use the strategy, as the prestigious agent cannot copy themselves. Pooling the estimate across all agents will cause the estimate to deflate towards 0 because in 1 of every 4 cases, the effect will be 0. A simple way to correct for this, is to add an interaction effect to the linear model, such that a different slope for prestige is assigned to prestigious and non-prestigious agents. To demonstrate this, shown below is the posterior distribution drawn from the population model for the prestige parameter pooled across the entire sample and two estimates which vary between the prestigious and non-prestigious agents.

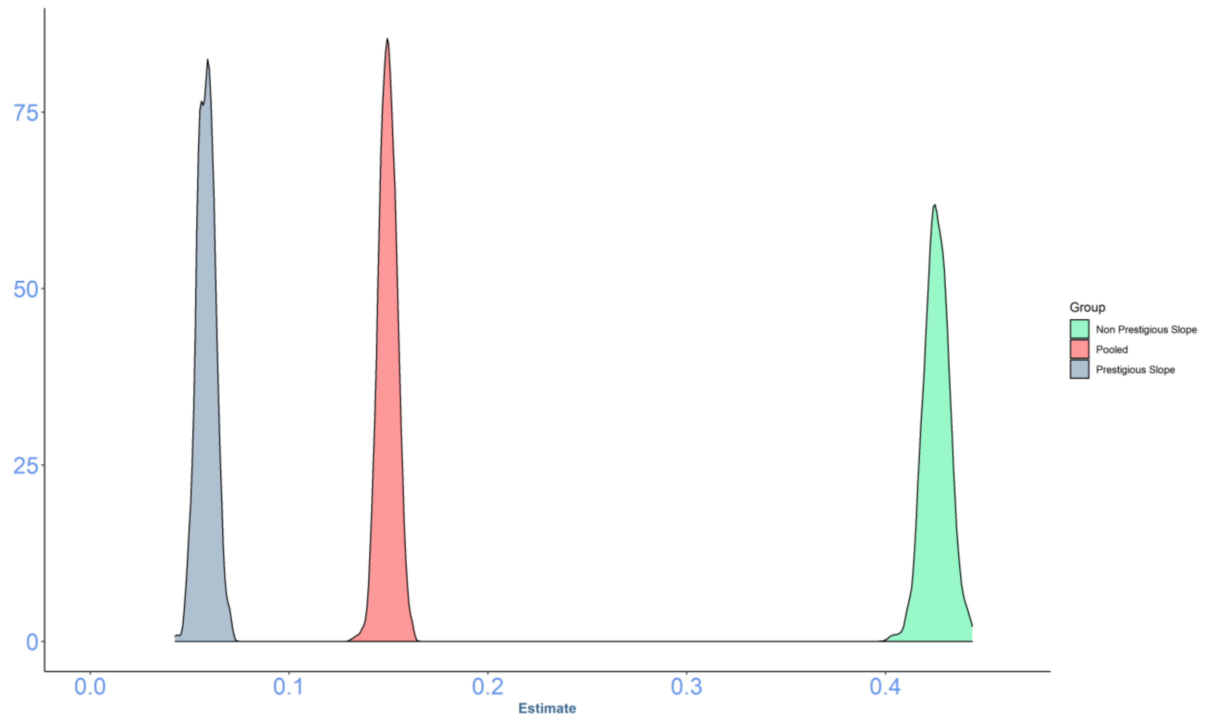


Figure 3.11*. Comparison between prestige-based learning slopes for prestigious (blue) and non-prestigious (green) against the pooled estimate (red). Plot shows 12000 random samples drawn from the posterior distribution of the population model.*

The pooled estimate falls directly between the two slopes which allow the interaction. The estimate for non-prestigious agents is now much larger and very reliably above 0, which is consistent with the known underlying mechanism for prestige in this model. Likewise, the estimated effect for being a prestigious agent is now far closer to 0. This supports the idea that allowing prestige to interact will increase the chance of detecting the effect. Otherwise, the confounded results in a deflated effect size, meaning it is impossible to distinguish between a true null effect and this statistical artifact.

This new model was therefore repeated on the sample sizes above to see if this improved detection of the prestige parameter. Plotted below in figure 3.12 is the slope association with prestige for just the non-prestigious agents along with the population estimate. The pattern is now closer to that of conformity and payoff learning. While uncertainty still exists in the model and the posterior means remain smaller than the true value, less of the samples overlap 0 and none of the posterior means are negative. The overall variation between the posterior means also seems to have reduced. Including the interaction facilitated greater detection of the true effect of prestige.

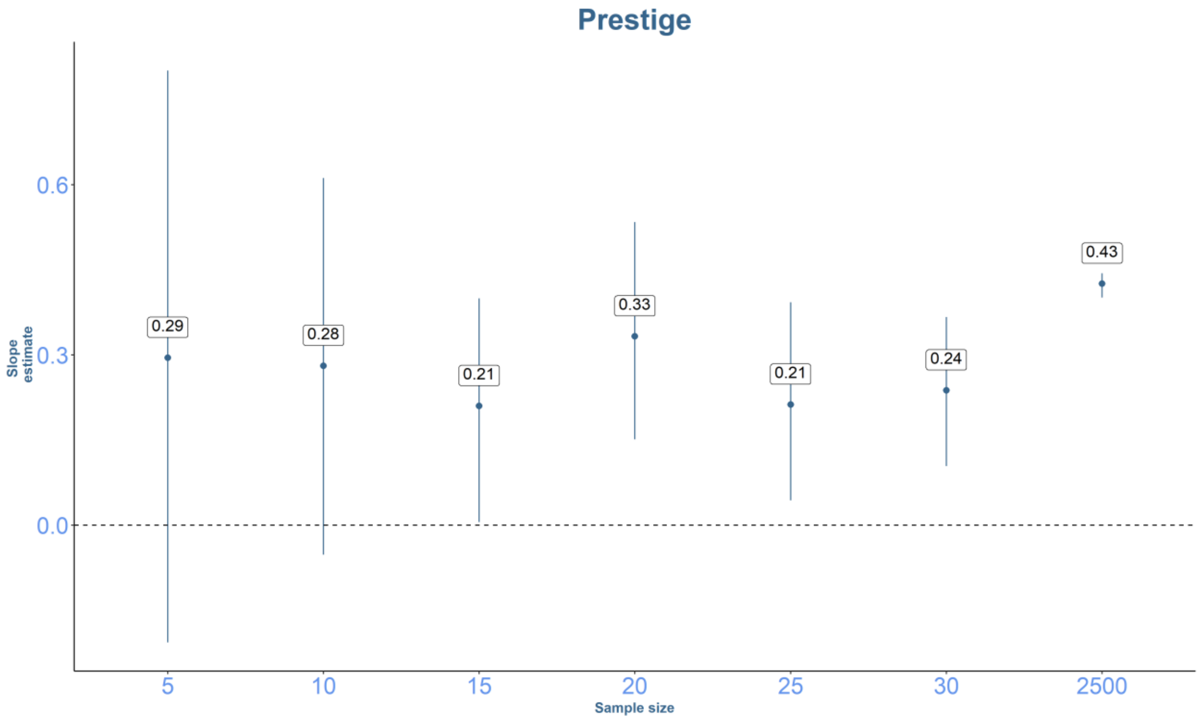


Figure 3.12. *Comparison of slope parameters for prestige from only the non-prestigious agent. Points show the mean value of the posterior distribution (labelled above) and the lines represent the minimum and maximum values of the whole distribution.*

Discussion

To help identify the required sample size for the laboratory experiment, simulations were run assuming that agents could modify their behaviour through conformity, prestige or payoff based social learning. Each round, agents sampled from a series of number to determine which (if any) strategy they would employ. Simulations were run with varying sample sizes and compared to a very large sample which represented the "population" or the "true" effect. Bayesian models were then fit to these simulations and the predicted parameter values associated with these strategies were extracted. The general trend observed in the population simulation was that cooperation trended towards 0 after 6 rounds, constituting a drop from its initial starting point.

Unsurprisingly, increasing the sample size increased the certainty of the models in every case. Conformity was the easiest for the model to identify, with the effect predicted to be reliably above 0 for samples above 15 and the mean being extremely close in every case except one outlier sample. Payoff learning was also well identified by the model above a sample of 15 though the associated parameter was less than conformity. The estimates values remained close to the true value but were less accurate than conformity. Prestige was more difficult for the model to detect: the mean estimate was still relatively close, but a greater number of samples overlapped 0 (and more considerably than the other strategies) and one sample erroneously predicted a negative coefficient. However, prestige still showed a reduction in uncertainty beyond a sample size of 10. This could be improved by introducing an interaction effect, which allowed the slope for prestige to vary between prestigious and non-prestigious participants. In this case, less of the samples overlapped 0 and the predicted effects were more similar across samples, suggesting it had been easier to detect.

Taken together, to detect the social learning strategies being used in a real experiment which would be subject to more random noise than was present here, it seems the minimum sample size should be 15 groups (60 participants). This will provide a good chance to detect an effect if one is present. In any case, one of the primary benefits of a Bayesian approach is that any uncertainty that the model has about the effect of any given parameter is explicitly plotted and reported (Kruschke, 2015; McElreath, 2020). If 15 groups of 4 participants turns out to be insufficient to detect the effect reliably, then this is the answer that will be provided by the model. This answer can then be used either to inform a further experiment or be used in simulation work which may be able to better highlight what the model is predicting. It should also be noted that this simulation applied only a limited range of possible parameter values (for example, it did not change the social learning probabilities). As such, this simulation is only intended to provide a practically useful ball-park estimate, rather than comprehensively explore the theoretical dynamics.

Appendix 4.1 – Free text responses

During debriefing and collection of demographic information, participants were asked to provide free text responses (of any length) to 2 questions. "*How do you think most people would say you should behave in this game?*" and, "*Why did you choose to behave how you did in the game?*". Overall, most responses were comments unrelated to the question such as "good" or were too vague to determine their meaning (for example, "like you normally would behave"). Where possible, comments were grouped according to themes in participant's responses.

Question 1 (Table 4.5) sought to identify if there was any pattern in social norms participants felt were associated with the game. Of those that could be grouped, 90 participants either directly suggested cooperative behaviour (or a synonym of this) or made some reference to "fair" or group beneficial behaviour. 21 suggested behaviours should be sensitive to the behaviour of others, which indicates concerns for reciprocity. In this sample, only 11 participants noted selfish behaviour as the expected norm. A further 14 participants indicated they didn't know or that they expected variability in individuals' opinions. Nevertheless, the general expected norm for participants in this experiment was for cooperative behaviour.

Table 4.5. *Summed responses from the question "How do you think most people would say you should behave in the game" grouped according to the theme of the response*

Theme	Count
Unrelated / Unclear comments	148
Behave cooperatively / alluding to "fair play"	90
Reciprocate others cooperation or behave tactically in response to others	21
Unsure / provided no strong indication	14

Behave selfishly or for your own personal benefit	11

Question 2 (Table 4.6) sought to better understand the conscious decision making made by participants in this experiment. The two most common themes (40 and 39 participants respectively) were comments suggesting maximising earnings or that decisions were based on observing their groupmates. It should be noted that, although 39 participants reported basing their decisions on their group mates, no participant made explicit mention of any social learning strategies. 26 participants claimed to be seeking to actively benefit their group and a further 9 did so irrespective of their group's contributions. This conforms somewhat with findings from experimental games which find most participants can be grouped into strong co-operators / free riders, or conditional co-operators (Kurzban & Houser, 2001). 17 participants referenced attempting to behave "fairly" to their group or that it was the "right thing to do". 5 participants reported following no particular strategy (reportedly choosing randomly, in some cases), 2 alluded to the warm glow effect (Andreoni, 1990) and feeling good from donating points and 1 participant reported wanting to "punish" those withholding points.

Table 4.6. *Summed responses from the question "Why did you choose to behave how you did in the game" grouped according to the theme of the response.*

Theme	**Count**
Unrelated / Unclear comments	141
Maximise Earnings	40
Influenced by group or levels of trust in their group	39
Benefit others	26
Fairness / morally right	17
Unconditional cooperation	9

No particular strategy	5

Overall, there was considerable variation in participants self-reported behavioural strategies but (aside from being generally influenced by their group) this did not include any explicit references to the social learning strategies addressed in this experiment. This ties into debates from the social learning literature regarding the conscious or unconscious use of social learning strategies (Heyes, 2016). In this case, the patterns associated with social information in our experiment do not appear to have occurred consciously, though the frequent mention of maximising earnings may imply the use of a payoff bias. Also of note is participants made no mention of the highest scoring participant in the quiz (the proxy for prestige), which further supports concerns that this individual was not granted high status by participants.

Appendix 4.2 – Conformity variable

As described in the main text, the conformity information used in the models included participant's own behaviour. Though this is consistent with what participants saw in the experiment, it means that participants contribute to their own social learning parameter, which may inflate the estimate associated with conformity. Therefore, the conformity information is reconstructed to exclude each participants behaviour from the average and the Prestige+Conformity+Payoff model was rerun with this corrected variable. The prestige and payoff bias variables remain the same. Figure 4.8 shows plotted model predictions from the new model for varying levels of social information and figure 4.9 shows model parameters from the corrected model plotted beside the Prestige+Conformity+Payoff model from the main text (uncorrected). The new conformity variable does not change the pattern of predictions. Payoff bias remains the strongest influence on participant cooperation while conformity and prestige remain weakly positive. The effect associated with conformity is virtually unchanged from that of the uncorrected model.

The parameter values are also virtually unchanged. The effect of the social condition remains negative while the difference between the SD and PD remains weakly positive. The social learning parameters did not change at all, meaning payoff bias remains as the most impactful social learning strategy. Therefore, the qualitative interpretations (and, largely, quantitative results) of the study remain unchanged whichever conformity variable is used in the analysis.

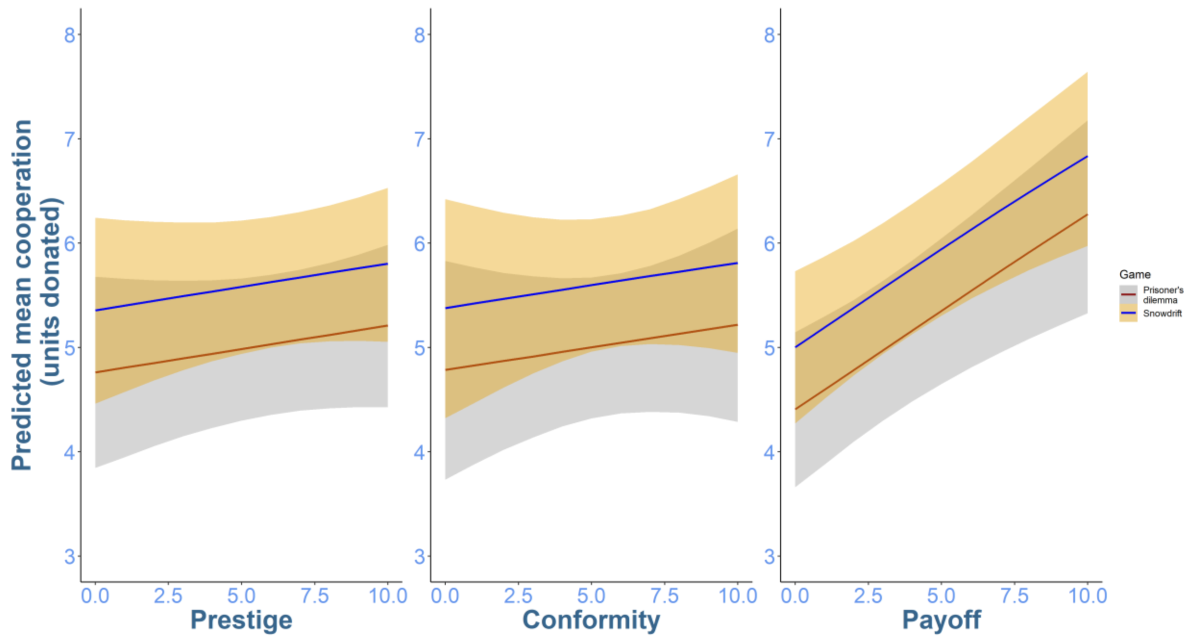


Figure 4.8. *Model predictions for the model with corrected conformity information varying the three social learning strategies and holding the other variables constant and setting the other social information at 0. Shaded regions represent the 95% prediction intervals for the mean. Prestige left; conformity middle; payoff; right*

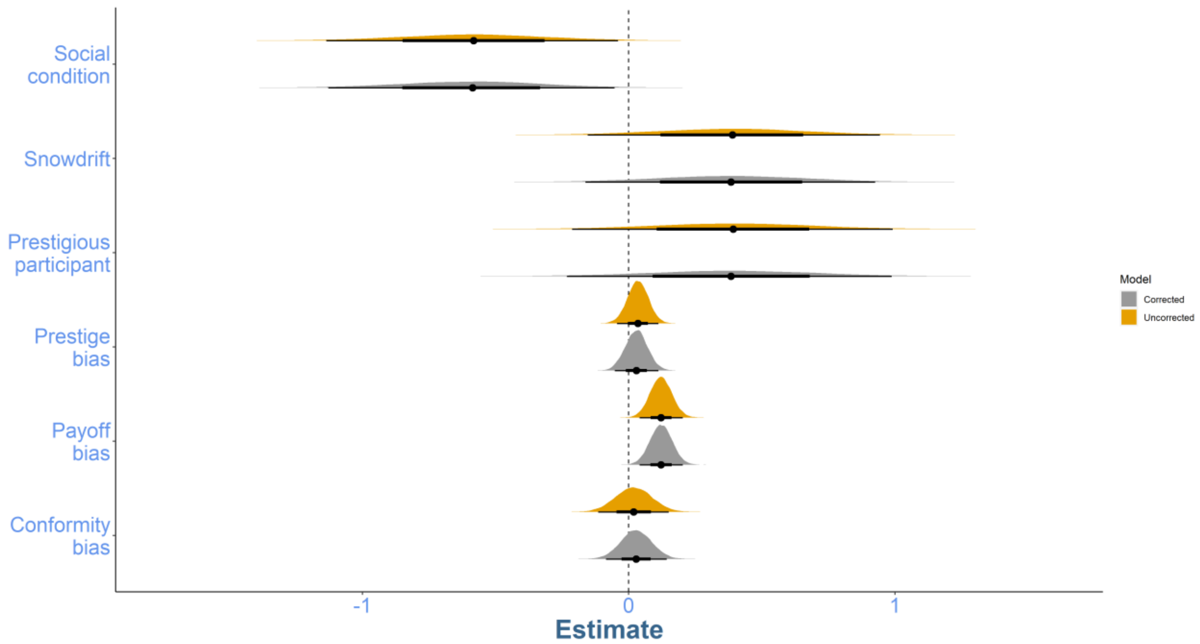


Figure 4.9. *Posterior distributions and 95% prediction intervals for all main effects in the Prestige+Conformity+Prestige model and the model for the uncorrected conformity information.*

Appendix 4.3 – Evaluating the impact of being prestigious on levels of cooperation

To evaluate the impact of being prestigious on cooperative behaviour, a similar strategy to that of evaluating social condition and game structure was taken. Note that the goal was to evaluate whether the subset of prestigious participants were more cooperative than non-prestigious participants. This is separate from the question addressed in the main text of whether participants were being influenced by the prestigious individual. The Prestige+Conformity+Payoff model from the main text was compared to an equivalent model which dropped the binary variable for prestige and another which also permitted an interaction with the social condition. This is to assess the possibility that prestigious individuals may have sought to lead their group by example, as is described in theoretical work (Henrich et al., 2015). If this is the case, it would be expected that increased cooperation would only occur in the social condition (where behaviour is observable). WAIC scores and model weights are reported below.

Table 4.7. *WAIC values and model weights for models evaluating the impact of prestige. Standard error difference shows the standard error in the difference between each model and the model with the lowest WAIC value.*

Model	*WAIC*	*Standard error difference*	*Weight*
Prestige * Social condition	5503.1	0	0.47
Prestige + Conformity + Payoff	5503.7	0.3	0.35
No Prestige effect	5505.1	0.6	0.18

As in the evaluation between game structure and social condition, there is similarly no clear distinction between any of the models. The small differences between WAIC indicate considerable uncertainty as to whether prestige should be in the model or not and that each model is likely to make very similar predictions. However, the models which include prestige are slightly favoured over the model that doesn't, which indicates some improvement in fit (combined weight = 0.82). To help visualise the predicted effects, model predictions from the Prestige * Social condition model are plotted below in figure 4.10. There is a slightly greater spread of predicted values for prestigious individuals which is likely explainable by the smaller number of observations. Otherwise, there is little indication of any difference between the prestigious and non-prestigious participants and further, no clear indication of any patterns in the interaction effect. The mean parameter estimate of the fixed effect for prestige in the interaction model was 0.41 (95% PI = -0.18; 1.00) which overlaps 0 quite considerably, indicating uncertainty in the effect. Overall, there is little evidence that prestigious individuals are more cooperative than non-prestigious individuals.

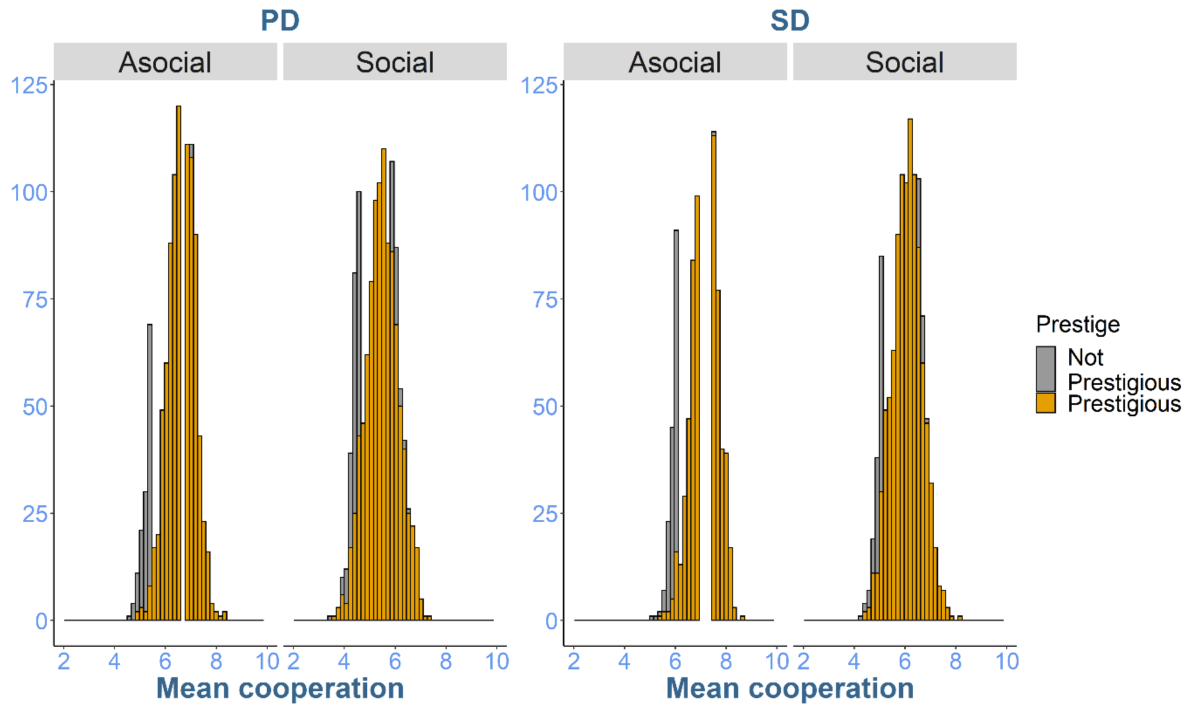


Figure 4.10. *1000 predictions of mean cooperation drawn from the posterior distribution across all experimental conditions for prestigious (grey) and non-prestigious individuals (yellow).*

Appendix 4.4 – Description of the simulation model

We constructed a simulation model which emulated the experiment and used the Bayesian posterior distributions and the Prestige+Conformity+Payoff model formula to affect changes in cooperation under various conditions.

At each round, each agent could contribute between 0 and 10 to the public good and then received a payoff calculated from the sum contributions of others divided by group size plus how much they kept (as in the experiment). Agents played in groups of varying sizes (group size is noted for each model below) across 100 rounds. This was repeated for 100 groups. One difference to the experiment was the prestigious agent was assumed to not partake in the public good. Instead, it was assumed that the prestigious agent could only be observed by the population, and their contribution level was set to either 2 or 8 (the value used is specified for each model). This kept the prestige effect consistent between groups as otherwise the prestigious agent's behaviour would change alongside the other agents. The population was finite and did not change size over time. For simplicity, we simulated only the PD and not the SD payoff structure used in the experiment. Agents change their behaviour using the model formula of the Prestige+Conformity+Payoff model shown below.

$$Cooperation \sim Categorical(\boldsymbol{p})$$

$$p_k = \begin{cases} q_k, & k = 1 \\ q_k - q_{k-1}, & k > 1 \end{cases}$$

$$logit(q_k) = \alpha_k - \phi$$

$$\phi = S + Intercept$$

$$S = (\beta_{social} + \beta_{prestige} * prestige\ bias + \beta_{payoff} * payoff\ bias + \beta_{conformity} * conformity)$$

At the start of the simulation, each agent samples from the posterior distribution intercepts (α_k) from the Bayesian model. These alongside the cooperation values of other agents in the current round are used to calculate *S*. The exception is the level of cooperation practiced by the prestigious agent, which is always set to 2 (low cooperation) or 8 (high cooperation).

At each round, we calculate a new cooperation value using the model formula above. A vector of probabilities for each level of cooperation (p), comprised of the probability of each individual level of cooperation (k), is used to generate a new value of cooperation from 0 - 10. p_k is calculated by reversing the logit on the vector of cumulative probabilities (q_k) which is calculated by subtracting the linear model term (ϕ) from the cut points (α_k) sampled from the posterior distribution. ϕ is given by the sum of *S* (social learning effect) and a varying intercept for each agent which can be thought of as their baseline propensity to cooperate irrespective of social learning. *S* is calculated using β values sampled from the posterior distribution alongside the social information of the current round. Note that to maintain the degree of parameter uncertainty expressed by the Bayesian model, each agent sampled unique values for the parameters each round when using horizontal transmission.

Order of events

Agents undergo the following steps. Steps 2 – 3 are repeated 100 times.

1) **Determine starting cooperation -** An agent's starting cooperation is determined by sampling from the distribution of cooperation observed in round 2 of the experiment. Additionally, each agent is also assigned an intercept value drawn from the posterior distribution.
2) **Public goods game -** Agents play in the public goods game and cooperate according to their stored level of cooperation. Like the experiment, the public good is multiplied by 2 and split between all individuals, regardless of contribution. An individual's payoff for the round is their earnings from the public good + their leftover contribution from 10.

3) **Modify behaviour –** Agents then independently modify their cooperation for the next round using the model formula shown above. Agents will copy the predicted level of cooperation exactly.

Results

See main text for main results.

Varying cooperative inclination of the prestigious individual

Here we compare dynamics across different constant contributions of the Prestigious individual, taking a value of 2 or 8. Figure 4.11 shows that the mean cooperation is weakly correlated with the degree of cooperation exhibited by the prestigious individual, but does not change the overall pattern of results

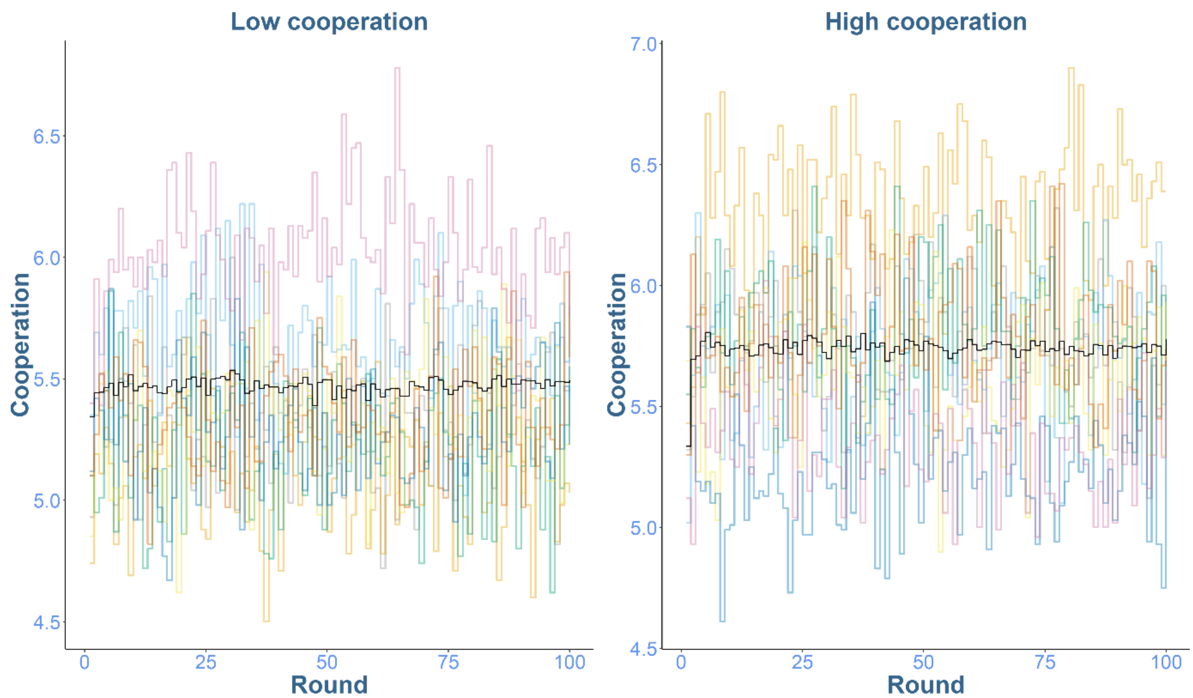


Figure 4.11. *Cooperation rates across 100 rounds (N = 100) where cooperation modifies by horizontal transmission only. Low cooperation from the prestigious individual (left) and high cooperation from the prestigious individual (right). Black line shows mean cooperation rates across all groups and coloured lines show cooperation rates for 8 randomly drawn groups.*

Varying group size

Here we compare the effect of group size upon mean levels of cooperation exhibited across the population. Figure 4.12 shows that mean cooperation rates stabilise to approximately the same level, irrespective of group size. Though, on average, smaller groups can sustain slightly higher cooperation levels than larger groups.

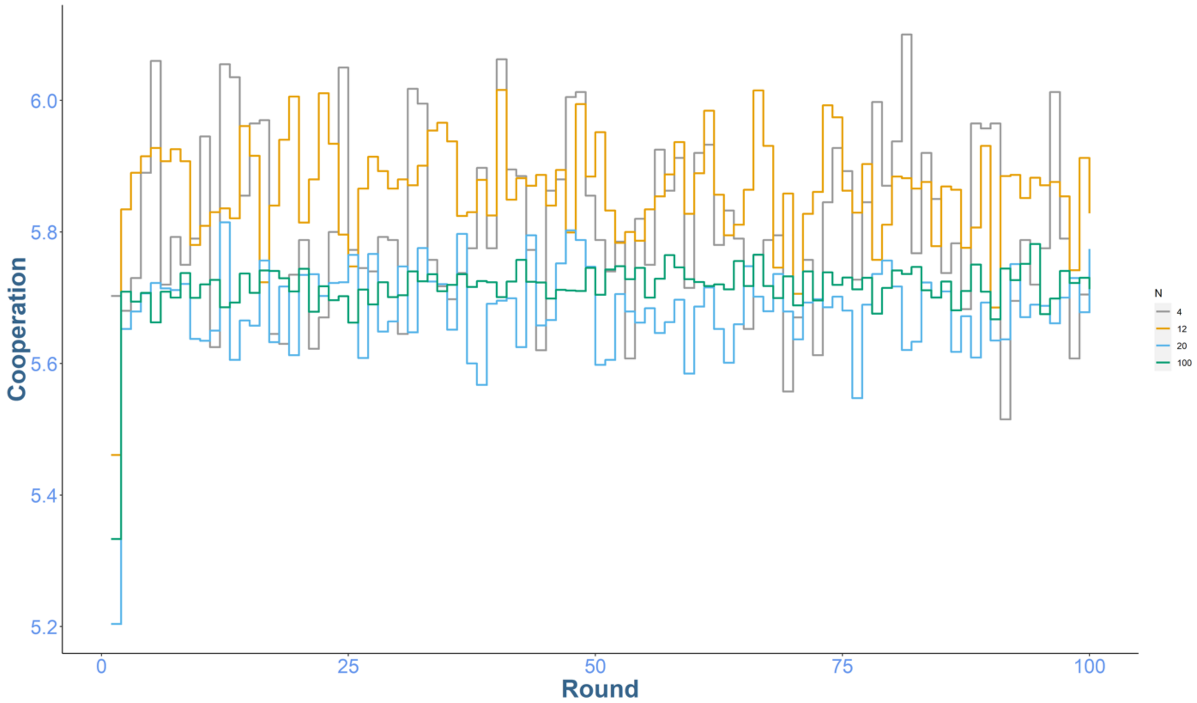


Figure 4.12. *Cooperation rates across 100 rounds where cooperation modifies by horizontal transmission only for varying group sizes (N = 4 (grey), 12 (gold), 20 (blue), 100 (green).*

Selection Model

The Selection Model assumes there is selection acting on intercept values in proportion to payoffs received, alongside mutation where a value drawn from a normal distribution with mean 0 and standard deviation 0.1 is added to each intercept.

In addition to results reported in the main text, here we report a condition where the initial distribution of intercept values is varied. A cooperative population was one where the population was comprised of intercepts sampled from values in the posterior distribution of above 0. An uncooperative population was one where intercepts were sampled from values in the posterior distribution of below 0. Higher intercept values are more inclined towards higher cooperation and vice versa. The results of this model across 1500 rounds is shown below in figure 4.13.

The overall result is unchanged from that of the main text. Both populations trend downwards from their initial starting point at similar rates. Cooperative populations sustain greater cooperation levels on average than uncooperative populations on account of the different starting intercept distributions.

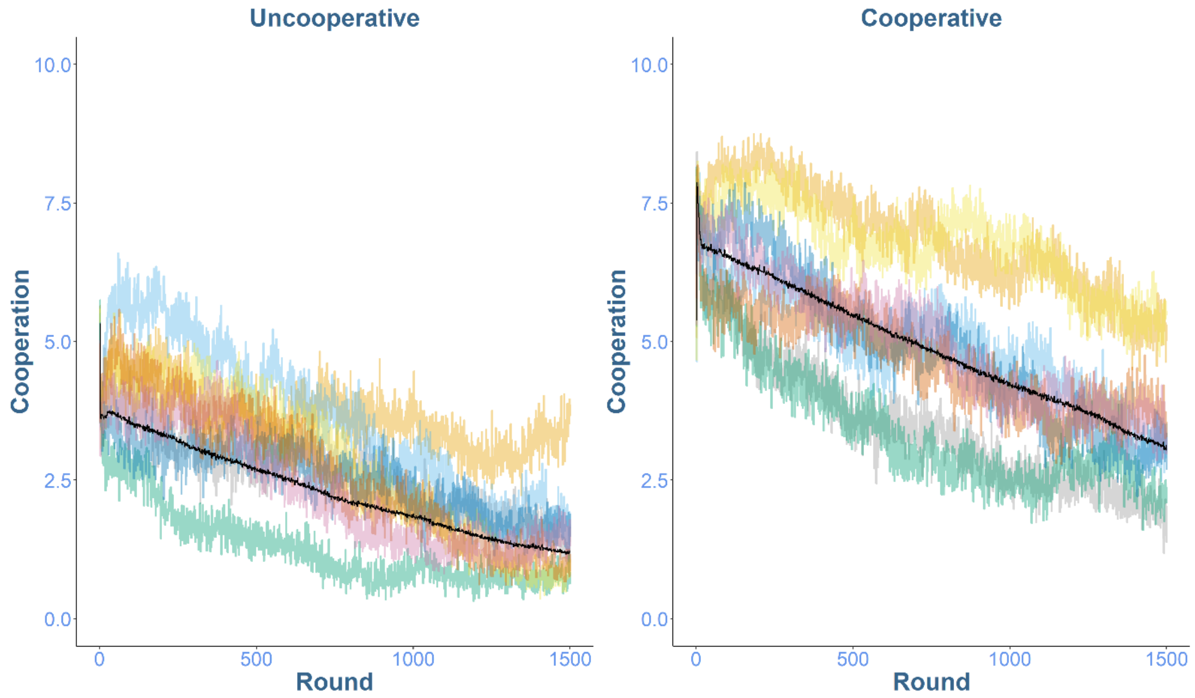
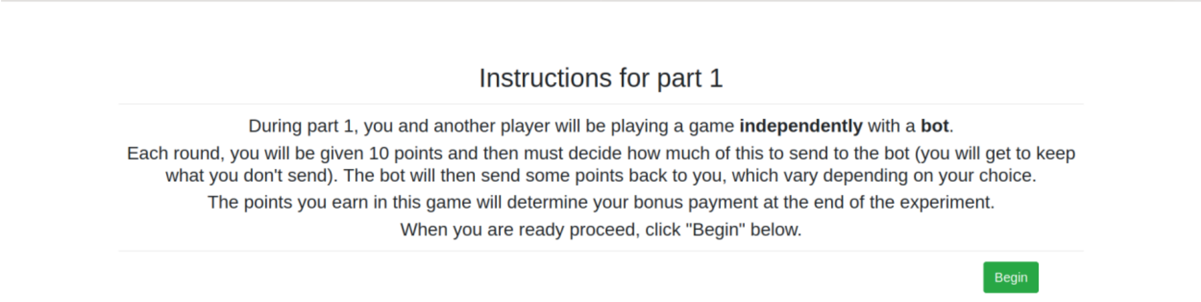
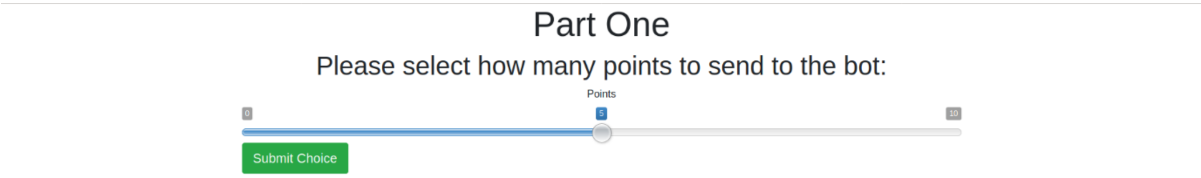


Figure 4.13. *Results from the selection model with varying starting compositions for the intercept values in the population between uncooperative (intercepts < 0, left) and cooperative (intercepts > 0, right). Black line shows mean cooperation rates across all groups and coloured lines show cooperation rates for 8 randomly drawn groups.*

Appendix 5.1 – Part 1 instructions and game interface





Appendix 5.2 – Part 2 instructions and interface

Part Two

You and the other participant (your partner) will complete the second part of the game **together.**

One of you will be assigned the **decider** role who can change the other participant's score.

After this, **the game is over**. The other participant **cannot** respond to their score being changed.

OK

Part Two

You have been assigned the **decider** role

Using a slider like the one below, you can **increase** your partners score by sliding to the **right** and **decrease** it by sliding it to the **left**.

Your score will decrease by 1 point for every 3 points you change your partner's score by.

Try sliding the slider below to see how this works.

To help you make your decision, you will be shown information about what **previous participants** chose to do as the decider.

Change to your partner's score:

0

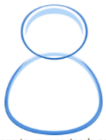
Your final score will be:	47
Your partner's score **will not change**	0

Next

Part Two

Will you pay points to change your partner's score? (remember, they cannot respond to your choice)

The highest scoring participant in previous games chose to:
Not change their partner's score



Change to your partner's score:

0

Your final score will be:	47
Your partner's score **will not change**	0

Submit Choice

Part Two

Please answer the following questions

Did you change your partner's score and if so, how?

Increased   Decreased   Did not change

Why did you make this decision?

Was your decision affected by the behaviour of the previous participants you were told about?

Submit responses

Appendix 5.3 – Qualitative responses for participants decisions

294 participants provided responses to the question asking them why they made their decision as the decider. 117 of these were either unrelated comments such as “VERY WELL” or were not clear enough to determine their meaning such as “It’s my opinion”. However, of the remaining 177 responses, it was possible for them to be grouped into themes (see table 5.4). 53 participants reported a desire to maximise their own points, which they could do by not doing anything. 52 participants alluded to a desire to help or benefit the other player or generally make sure that they also had some points (which did indicate some misunderstanding of the experimental setup in some cases). A further 19 participant made direct reference to giving points being “fair” or that their decisions were to make the points equal between participants (which again, indicated some misunderstanding). 9 participants explicitly wanted the points to remain the same and 7 reported their decision was at random. The remaining 37 participants offered many other miscellaneous reasons which did not fit neatly into these categories including wanting to harm their partner, donating points being efficient (because of the 1 – 3 ratio) or 1 participant that specifically wanted their score to be 40.

Table 5.10, *Breakdown of the counts of participant responses to the question “Why did you make this decision” about their choice as the decider*

Theme	**Count**
Unrelated / unclear comments	117
Want to maximise own points	53
Want to share / benefit others	52
Wanted points to be equal / fair	19
Wanted things to remain the same	9

Decision was random	7
Another reason	37

Appendix 6.1 – Vignette text

Conformity text:

A recent UK survey commissioned by gocompare.com has found that 70% of the UK population agree that greenhouse gases are to blame for the rise in temperature. Many businesses and citizens are now making increased efforts to reduce waste and curb emissions to protect the environment. Alex's neighbours have just bought an electric car and many people on social media have been recommending the plastic free supermarket. Alex's neighbourhood is now one of the greenest in the area. Inspired by this, many other neighbourhoods are now also reducing their emissions.

Payoff text:

It was since Alex started making a conscious effort to behave more sustainably a few years back that they began to notice the benefits. Avoiding using the heating during winter saved Alex a considerable amount of money in bills and it was intensely satisfying to see the amount of waste that could be avoided by spending that little extra time separating rubbish and remembering to bring shopping bags. People at work have also noticed what Alex does for the environment and have expressed their admiration and respect.

Prestige text:

Alex was a climate advocate long before their award-winning performances on stage and screen. "Climate change is real, it is happening right now, it is the most urgent threat facing our entire species, and we need to work collectively together and stop procrastinating" the actor said. Throughout their career, Alex has donated hours of time, vast amounts of funding and discretely attended climate marches to help advance United Nations climate negotiations, protect coral reefs and tigers and spread public awareness.

Note: This text is paraphrased from a quote by Leonardo DiCaprio

Appendix 6.2 – Free text responses

Only 32 participants offered anything in the free text response box, and many were themed around items included in the questionnaire. Several comments also fit multiple categories, so were grouped in each one that was appropriate (giving 35 comments total). The most common category of responses (15) were comments elaborating on reasons that sustainability is either unavoidable or unclear. For example, not being able to control the amount of packaging on online orders or being unable to use more sustainable alternatives due to disability. 6 comments referenced sustainability being the responsibility of large corporations and 2 that being sustainable was pointless as an individual. 3 comments cited lack of time or sustainability being inconvenient and 2 that they lacked the necessary knowledge. 2 comments expressed concern that others may find them "preachy" if they were to point out others unsustainability and 2 others that they did not want to give us the less sustainable alternatives. 1 comment highlighted a particular hygiene concern with washing clothes at lower temperatures. The remaining 2 were unclear and could not be grouped into any category. Overall, the most common theme were comments pertaining to systematic barriers, which is broadly consistent with the opinion expressed by the interview participants.

Appendix 6.3 – Descriptives and correlations between factor scores and predictors included in the Bayesian models

Table 6.9. *Mean and standard deviation of (standardised) factor scores between females and non-females and vegans and non-vegans.*

		Perceived Duty	Perceived Barriers	Acting sustainably
Non-female	Non-vegan	-0.16 (1.18)	0.06 (1.06)	-0.35 (1.26)
Non-female	Vegan	0.02 (0.76)	-0.27 (1.17)	-0.34 (1.08)
Female	Non-vegan	-0.12 (0.98)	0.16 (0.96)	0.02 (0.98)
Female	Vegan	0.39 (0.80)	-0.26 (0.94)	0.18 (0.91)

Table 6.10. *Correlation matrix showing correlation coefficients (r) between all continuous variables included in the Bayesian models.*

	Conservatism	**Perceived duty**	**Perceived barrier**	**Acting sustainably**	**Durham University student norms**	**Family Norms**	**Friends' norms**
Conservatism	1.00	-0.38	0.07	-0.03	-0.01	0.08	-0.08
Perceived duty		1.00	-0.24	0.23	-0.02	0.05	0.18
Perceived barrier			1.00	-0.27	-0.12	-0.23	-0.17
Acting sustainably				1.00	0.03	0.09	0.08
Durham University student norms					1.00	0.11	0.39
Family Norms						1.00	0.25
Friends' norms							1.00