EMPIRICAL INVESTIGATIONS OF SOCIAL LEARNING, COOPERATION, AND THEIR ROLE IN THE EVOLUTION OF COMPEX CULTURE

Cara L. Evans

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

2016

EMPIRICAL INVESTIGATIONS OF SOCIAL LEARNING, COOPERATION, AND THEIR ROLE IN THE EVOLUTION OF COMPLEX CULTURE

Cara L. Evans

University of
St Andrews

University of St Andrews

Candidate's declaration:

I, Cara Evans, hereby certify that this thesis, which is approximately 58,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in April 2012 and as a candidate for the degree of Ph.D. in April 2016; the higher study for which this is a record was carried out in the University of St Andrews between 2012 and 2016.

Date: 22/04/16 Signature of candidate:

Supervisor's declaration:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Ph.D. in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date: 22/04/16 Signature of supervisor:

Permission for publication

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the print and electronic publication of this thesis:

No embargo on print copy. No embargo on electronic copy.

Date: 22/04/16 Signature of candidate:

Signature of supervisor:

Contributions

Chapter 3

Dr. Rachel Kendal provided feedback on the experimental design and interpretation of the results. Prof. Malinda Carpenter provided advice on the interpretation of the results. Shilpa Jujjavarapu double-coded video data for use in tests of inter-observer reliability. All other work was carried out by myself.

Chapter 4

Shilpa Jujjavarapu double-coded video data for use in tests of inter-observer reliability. Dr. Michal Arbilly, Camille Troisi and Riva Riley acted as experimental confederates. All other work was carried out by myself.

Chapter 5

Dr. Tom Morgan provided assistance with the implementation of custom-made experimental computer games in Java. All other work was carried out by myself.

Acknowledgements

I am privileged to have spent the last few years here in St Andrews. It really was the best place for me to undertake a PhD, and for that I will always be indebted to my supervisor, Kevin Laland. Thank you for your support, your kindness, for re-introducing me to whisky after I had long given up on it, and for being lots of fun.

I am also very grateful to Rachel Kendal. Rachel's enthusiasm for discussing science, her warm welcome and encouragement, are what began my journey here to St Andrews, and I am thankful for her continued advice. Malinda Carpenter has provided invaluable guidance and feedback on the work presented in Chapter 3. I have already learned so much from Malinda, despite only having recently begun working with her.

Within the St Andrews community, it has been a pleasure to meet many wonderful colleagues and friends – too many for me to thank each individually. I was particularly fortunate to belong to a great group of office mates, both in the Bute building (never forgotten!) and then later in the HMB, whom I especially wish to thank: To Laura, Laurel, Tom, Alice, Camille, Riva, Sally, Andrew, James, Elena and Murillo – thank you for putting up with me! It wouldn't have been the same here without you.

Many thanks also go to Michal Arbilly, Camille Troisi and Riva Riley, who, in giving up way too much of their time to help with experiments, were willing to reveal their secret talent for acting. I am also indebted to Tom Morgan, who takes all credit for ensuring my brain was not lost somewhere on the Java 'heap'.

The work presented in this thesis was supported by an ERC grant awarded to Kevin Laland and by a BBSRC studentship awarded to me. I am also grateful to Katherine Meacham, who helped prevent many an administrative disaster.

We are all, of course, the product of our history, and I am thankful to many great family members, teachers, and friends who have contributed to mine. Especially to Gill, Rob, Jack, Mavis, Denzil, and Tronnie (best cat ever) – I was lucky to have you on my journey.

General Abstract

There is something unique about human culture. Its complex technologies, customs, institutions, symbolisms and norms, which are shared and maintained and improved across countless generations, are what sets it apart from the 'cultures' of other animals. The fundamental question that researchers are only just beginning to unravel is: How do we account for the gap between their 'cultures' and ours? The answer lies in a deeper understanding of culture's complex constituent components: from the micro-level psychological mechanisms that guide and facilitate accurate social learning, to the macro-level cultural processes that unfold within large-scale cooperative groups. This thesis attempts to contribute to two broad themes that are of relevance to this question. The first theme involves the evolution of accurate and high-fidelity cultural transmission. In Chapter 2, a meta-analysis conducted across primate social learning studies finds support for the common assumption that imitative and/or emulative learning mechanisms are required for the high-fidelity transmission of complex instrumental cultural goals. Chapter 3, adopting an experimental study with young children, then questions the claim that mechanisms of high-fidelity copying have reached such heights in our own species that they will even lead us to blindly copy irrelevant, and potentially costly, information. The second theme involves investigations of the mutually reinforcing relationship predicted between cultural complexity and ultra-cooperativeness in humans, employing a series of laboratory-based experimental investigations with adults. Chapter 4 finds only limited support for a positive relationship between cooperative behaviour and behavioural imitation, which is believed to facilitate cultural group cohesion. Finally, Chapter 5 presents evidence suggesting that access to cultural information is positively associated with an individual's cooperative reputation, and argues that this

dynamic might help to scaffold the evolution of increased cultural complexity and cooperation in a learning environment where cultural information carries high value.

Table of Contents

Chapter 1

General Introduction

1.1. Human culture and animal 'culture'

Culture sets human beings apart from the rest of the animal kingdom. Because culture is underpinned by social learning, human cultural processes are able to modify cognition and behaviour, bring about environmental change, and accumulate increasingly complex modifications, over a far shorter timeframe than genetic or individual-driven change (Laland & Janik, 2006; Thompson et al., 2016; Tomasello, 1999). *Social learning*, broadly defined, refers to 'learning that is facilitated by observation of, or interaction with, another individual, or its products' (Hoppitt & Laland, 2013, p4). Behavioural patterns implicating social learning have also been identified across a range of non-human animals – notably fish, cetaceans, primates, birds and insects – but there is little consensus regarding which, or what, if any, socially transmitted behaviours in other species should be regarded as culture (Laland & Hoppitt, 2003; Price et al., 2010; Leadbeater & Chittka, 2007).

The relative abundance or scarcity of culture in nature ultimately depends on where we etch culture's boundaries. At one extreme, culture becomes the exclusive domain of *Homo sapiens*, replete with language, teaching and symbolism (e.g., Tuttle, 2001). At the other, culture might be any inherited group level differences in behaviour that are not attributable to genetic variation; a definition that encompasses a wide range of species, including some bacteria (Lumsden & Wilson, 1981). Throughout this thesis I employ a broad and inclusive definition of *culture*, requiring only that 'group-typical patterns of behaviour' exhibit evidence of 'socially learned or transmitted information' to be considered cultural (Hoppitt & Laland, 2013, p4).

Whether or not we allow other species the accolade of 'culture', many evolutionary minded scientists now accept that the so-called 'proto-cultures' or 'behavioural traditions' of great apes, some species of monkey, cetaceans and birds

exhibit some semblance to human culture (Whiten et al., 1999; van Schaik et al., 2003; Laland & Janik, 2006). Assuming the broad and inclusive definition of culture deployed above, there is now good evidence for culture in the chimpanzee (Hobaiter et al., 2014; see Lonsdorf, 2006, for further support), humpback whales (Allen et al., 2013; Garland et al., 2011; Noad et al., 2000), at least two species of birds (Slagsvold & Wiebe, 2007; see also Aplin et al., 2015), and fishes (Helfman & Schultz, 1984; Warner, 1988). It is likely that a growing list of taxa that can be said to exhibit 'cultural' behaviour will soon emerge, as a suite of new and innovative methodologies designed to detect social learning in naturalistic contexts, are increasingly applied (see Hoppitt & Laland, 2013).

What is not debated among interested scholars, is the gulf in complexity and magnitude that exists between the cultural worlds of other species and of our own (Laland & Hoppitt, 2003). Cultural capacities unique to humans have driven unprecedented population expansions into diverse environments, and the establishment of highly cooperative societies scaffolded by complex social institutions, such as religions and central governments. We alone appear to have stepped up from the technologically stagnant and often fleeting behavioural traditions typically characterised in other species (Laland et al., 1993; Thornton & Malapert, 2009), accumulating cultural improvements, generation on generation, that have allowed us to write books, operate machinery and drive cars. There have been claims of cumulative culture in a handful of animal species, most notably chimpanzees and New Caledonian crows (Boesch, 2003; Hunt & Gray, 2003; Yamamoto et al., 2013). The evidence for these claims is circumstantial and equivocal, however, and the purported culturally accumulated behaviours still rudimentary next to our own (Dean et al., 2014; Tomasello & Herrmann, 2010).

We have to acknowledge that the cultural evolutionary process has reached unprecedented levels in our own relative to other species. The fundamental question we are now left with is: How do we account for the gap between them and ourselves? This might be best viewed as a question with two parts (Rendell et al., 2011a). Firstly, we must ask: What specific properties of human social learning and interaction give rise to the complexities found in our own but not other animal cultures? Then later: How did the processes that set human culture apart evolve out of the types of more rudimentary abilities possessed by other animals?

The body of work presented in this thesis aims to contribute to the first of these questions. Traditional approaches have sought to elucidate human cultural uniqueness in specific cognitive capacities owing to our exceptionally large brains. In particular, research efforts have focused on identifying the behavioural and psychological learning processes responsible for high-fidelity cultural transmission between individuals (e.g., Boyd & Richerson, 1985; Tomasello, 1999). **Chapter 2** and **Chapter 3** attempt to further our knowledge of the learning mechanisms and decision-making processes involved in high-fidelity social transmission. Increasingly, however, theoretical and empirical approaches are highlighting the importance of macro-level processes in cultural transmission – such as demography and large-scale interaction – and the interplay between macro-level processes and cultural cognition in fashioning cultural complexity (e.g., Cantor & Whitehead, 2013; Hill et al., 2011; Powell et al., 2009; Pradhan et al., 2012). **Chapter 4** and **Chapter 5** attempt to examine the interplay between the mechanisms underpinning group level cooperation and the mechanisms involved in cultural transmission.

Before turning to these projects in more detail, it is first necessary to provide the reader with an overview of some key topics and definitions pertaining to the work

presented. These include social learning mechanisms; high-fidelity social learning; social learning strategies; and the evolution of large-scale human interaction, cooperation and culture.

1.2. Social learning mechanisms

A number of social learning taxonomies have emerged in recent years, which have attempted to describe the different psychological processes through which one animal might learn from another (Galef, 1988; Heyes, 1994; Whiten & Ham, 1992; Whiten et al., 2004; Zentall, 2001). These taxonomies have only offered partial agreement on terminology and definitions, however, and have been criticised for assigning the proposed learning mechanisms to different levels of cognitive complexity based on subjective assessments (Roitblat, 1998). Hoppitt and Laland (2013) describe 12 social learning mechanisms that can each be distinguished empirically, yet concede that limitations present in much of the currently available data on animal social learning largely obscure accurate categorisations. The operation of more than one mechanism at any given time is also a possibility (Hoppitt et al., 2012).

Human beings are regarded as capable of utilising all identified learning mechanisms, yet it has been hotly debated whether non-human animals (particularly non-human primates) are capable of learning via mechanisms considered to be more cognitively enhanced (Heyes, 2012b; Tennie et al., 2009; Whiten et al., 2009). Here I will outline four categories of learning mechanisms – imitation, emulation, stimulus and local enhancement – that are of most relevance to the body of work presented in

this thesis, and refer the reader to Hoppitt and Laland (2013) for a more exhaustive overview.

Imitation

Imitation is the most contentious and widely discussed mechanism within social learning research, regarded by many as the most cognitively complex form of social transmission (Hoppitt & Laland, 2008), and by some researchers as a capability unique to humans (e.g., Tomasello, 1999). During the last few decades, the definition of imitation within the field of social learning has been subject to much debate (Whiten et al., 2004). For example, Whiten and Ham (1992) described imitation broadly, as a process by which "B learns some aspect(s) of the intrinsic form of an act from A" (p.250). Other authors have proposed narrower criteria, insisting on a distinction between *imitation*, which they define as copying the action topography of an agent's body movements, and *emulation* (see below), which refers to copying the environmental effects of the agent's actions (Tomasello & Call, 1997; Tomasello, 1998). Other debates have concerned whether imitation of another's actions must incorporate an understanding of their intentional states and goals (Heyes, 1998; Horowitz, 2003; Tomasello, 1999), and whether the particular action copied must be novel to the learner (Byrne & Russon, 1998; Byrne, 2002).

Throughout this thesis, I use the term *imitation* to refer to evidence that an observer has copied some part of the body action topography, or action sequence, of another individual. A distinction is acknowledged between *simple* or 'mindless' forms of imitation and *complex* imitation, which is believed to be more controlled, cognitive, and involved in the acquisition of novel behaviour (Heyes, 2011; Tomasello, 1996). The latter, more complex form, regarded by some as exclusive to human beings (e.g.,

Tennie et al., 2009), is frequently the focus in experimental studies of instrumental social learning (although see Byrne, 2002, for a contrary view), and is most relevant to discussion in **Chapter 2** and **Chapter 3**. The former, simpler form, also referred to as automatic imitation (Heyes, 2011), nonconscious mimicry (Chartrand & Bargh, 1999), contagion (Thorpe, 1963), response facilitation (Byrne, 2002), or social facilitation (Visalberghi & Addessi, 2000), and regarded as taxonomically more general, involves the unintentional imitation of a motor pattern that was already part of the imitator's behavioural repertoire. Simple imitation is widely implicated in processes such as behavioural contagion, behavioural synchrony and group cohesion (Chartrand & Van Baaren, 2009), and forms the basis of discussion in **Chapter 4**. However, while *simple* or *complex* forms of imitation are often inferred, empirically it might often be difficult to tease apart one form from another (Byrne, 2002; Heyes & Ray, 2000), and a recent experimental investigation suggests that both processes might often be operating simultaneously (Belot et al., 2013).

Emulation

Emulation has also been subject to different definitions within the field of social learning (Byrne, 2002; Whiten et al., 2004), but can broadly be considered as describing the process by which an animal copies the environmental results of another's actions, rather than the bodily actions themselves (Hoppitt & Laland, 2013, p77). Thus, an example of emulation might be an observer tying to recreate the object movements made by another individual (e.g., making a door move in the same manner/direction as another moved it). Alternatively, emulation could also constitute an observer trying to recreate the end product of another's behaviour (e.g., an artefact

that another has created), without use of knowledge of the action or object movements that were involved in the product's initial creation.

Stimulus and local enhancement

Stimulus and local enhancement are commonly regarded as abundant throughout the animal kingdom and are considered two of the simplest forms of social learning (Spence, 1937; Thorpe, 1963). Stimulus enhancement is said to occur when an individual's behaviour results in a second observer individual becoming more responsive, and more likely to interact with, a single stimulus (Hoppitt & Laland, 2013). For example, an individual performing the target behaviour on a stimulus, through increasing a second observer individual's interaction with the same stimulus, might subsequently lead to an increase in the rate at which the observer learns and acquires the behaviour performed by the first individual. Similarly, local enhancement occurs when exposure to an individual's presence, or interaction with objects, at a particular location, results in an increase in the rate at which an observing individual visits, or interacts with objects, at that same location (Hoppitt & Laland, 2013).

Identifying social learning mechanisms

Despite on-going definitional debates in imitation research, the majority of scholars have generally agreed upon the adoption of the two-action task for assessing imitative learning abilities in instrumental contexts (Dawson & Foss, 1965; Galef, 1988; Whiten & Ham, 1992; Heyes, 1993; Heyes & Ray, 2000; Zentall, 2001; for a contrary view see Byrne, 2002). Developed by Dawson and Foss (1965), the two-action task requires all test subjects to solve a task with two alternate solutions, but each subject is randomly assigned to witness only one solution. If the group of subjects show a

disproportionate tendency to match the solution they observed, imitation is assumed; otherwise they are assumed to have been reliant on alternative learning mechanisms, such as local or stimulus enhancement. However, evidence of imitative learning can only be confirmed if all other possible learning mechanisms can be ruled out. Heyes and Ray (2000) duly noted that most applications of the two-action task do not discriminate between *imitation* (copying the action topography of the demonstrator's body movements), and *emulation* (reproducing the task-related object movements that resulted from the demonstrator's actions).

Some researchers, in attempts to disentangle the relative importance of imitative versus emulative learning, have assigned test groups to conditions where the apparatus was moved surreptitiously by the experimenter (dubbed "ghost" conditions by Fawcett et al., 2002), denying subjects information about the motor actions required to operate it (see Hopper, 2010). The reverse of the ghost condition, where subjects are shown the body actions necessary to operate the apparatus, but not the accompanying object movements, has also been implemented (Call et al., 2005). Results generated through the use of ghost conditions should be interpreted with caution, however, as the non-social context in which the demonstration occurs might be intrinsically less appealing to animals than contexts where another individual is present (Byrne, 2002).

While categorisations of social learning mechanisms are explicit in their descriptions of the *behavioural* processes through which an animal might socially learn, the *neural* underpinnings of each mechanism are still poorly understood. This has obvious implications for existing assumptions concerning the complexity of one learning mechanism relative to another, which might prove to be incorrect, and also for claims regarding the evolved specificity of mechanisms for use in social learning

(Heyes, 2012b). A number of neural mechanisms that might underpin social learning, especially imitative learning, have been proposed, but will not be discussed further here (see Brass & Heyes, 2005; Heyes, 2009; Hoppitt & Laland, 2013).

1.3. High-fidelity social transmission

The fidelity of social transmission – that is the accuracy and efficiency with which learned behaviours can spread between individuals – is considered pivotal to the remarkable cultural achievements of our species (e.g., Boyd & Richerson, 1996; Galef, 1992; Lewis & Laland, 2012; Morgan et al., 2015a; Tomasello, 1999). It has been argued compellingly that the faithful and efficient transfer of culturally learned information is what allows each new generation of learners to build upon the achievements of its predecessors: a process that has been termed 'cultural ratcheting' (Tomasello et al., 1993). By this view, human beings have evolved the high-fidelity learning mechanisms necessary to initiate and fuel the cultural ratcheting process, while other animals, reliant on more erroneous, lower-fidelity forms of social transmission, must reinvent the wheel each time (Tennie et al., 2009).

Support for this argument is offered by a recent comparative study in which a battery of cognitive tests was conducted across groups of chimpanzees, orang-utans and 2.5-year-old children (Herrmann et al., 2007). The results of this study suggested that cultural rather than general intelligence appears to be accounting for the human–nonhuman cultural gap: while each of the species exhibited rather similar abilities across cognitive domains associated with the physical world (i.e., assessing quantities, understanding space and causality), the human children far exceeded the other ape

species on skills relating to the social world (i.e., social learning, communication, and shared intentionality; but see de Waal et al., 2008, for a critique of these findings).

Of course, social learning must still work in tandem with novel invention and causal reasoning in order to accumulate cultural improvements: high-fidelity transmission and innovation are thought to be the dual engines driving cultural complexity. To this end, recent experimental evidence suggests that human cultural prowess may result from an ability to apply high-fidelity copying selectively and strategically in combination with skills such as individual learning (innovation) and trait recombination (Derex & Boyd, 2016; Legare & Nielsen, 2015; Wood et al., 2015; Zwirner & Thornton, 2015; for further discussion on selective social learning see 'Social learning strategies' below and **Chapter 3**). This suggestion chimes with other empirical evidence implying that, across primate species at least, social and individual learning abilities have coevolved (Reader & Laland, 2002; Reader et al., 2011).

Models and mechanisms of high-fidelity learning

The case for fidelity as a key player in the evolution of human cultural complexity has received support from a number of theoretical models. For example, using a series of analytical models to investigate the factors underpinning cultural stability and transmission, Enquist et al. (2010) found that, above a given population threshold, relatively small increases in transmission fidelity could lead to disproportionately large increases in both the amount and longevity of cultural traits. These findings suggest that, through ensuring the perseverance of cultural traits for long enough to undergo cumulative modification, just small shifts in transmission fidelity might have been enough to initiate and fuel the cultural ratcheting that now typifies our species.

Likewise, Lewis and Laland (2012), employing simple mathematical models of cultural trait accumulation, discovered that fidelity (here represented by the inverse of a parameter that determined the relative rate at which useful traits could be lost from the population) was far more deterministic of the level of cultural complexity achieved than other contributory factors, such as new innovations, modifications, or the recombination of existing traits. When fidelity was below a certain threshold, cumulative culture did not get off the ground.

Transmission fidelity in theoretical investigations of cultural evolution has been modelled explicitly as a function of learning outcomes (i.e., the frequency, complexity, perseverance or the payoffs associated with cultural traits), adopting the 'phenotypic gambit' and bypassing consideration of the particular psychological and behavioural mechanisms that give rise to faithful transmission (e.g., Enquist et al., 2010; Lewis & Laland, 2012; Rendell et al., 2010a). The particular psychological and behavioural mechanisms that might afford cultural fidelity have instead been the domain of psychologists and biologists interested in comparative cultural cognition. The relative importance of the different social learning mechanisms has been much debated. Imitation has received a special status (e.g., Galef, 1992; Tennie et al., 2009; Tomasello, 1999), yet the evidence for an especial role of imitative learning in instrumental cultural transmission is equivocal (see below).

Is imitation required for instrumental learning and cultural ratcheting?

The elevated status awarded to imitation in explanations of human cultural complexity is questionable in light of evidence that nonhuman animals also engage in complex imitation. Marmosets are capable of imitating the actions used by a demonstrator to remove a lid from a canister with remarkable fidelity, for instance (Voelkl & Huber,

2007). Likewise birds, such as pigeons (Saggerson et al., 2005) and quail (Akins & Zentall, 1998; Dorrance & Zentall, 2001), will imitate a conspecific's actions if those actions have previously led to a desirable food reward. While this does not preclude the possibility that human imitation is quantitatively different from imitation in other species – humans may be capable of imitating more detailed or more complex sequences of actions – it does suggest that human imitation need not result from qualitatively dissimilar neural underpinnings (Heyes, 2012b). However, that imitation is found in animals taxonomically distant from humans suggests that imitation per se does not account for the human-nonhuman cultural gap. Indeed, even those researchers who emphasise imitation in explanations of human cultural uniqueness, regard it as necessary, but not sufficient, for cumulative culture (Tennie et al., 2009; Tomasello, 1999).

Yet testing for evidence of imitation in other species is itself a qualitatively different question than asking: Does imitation lead to increases in the fidelity (i.e., success/accuracy) of cultural trait transmission? Or, Is imitation necessary for cultural ratcheting? Direct attempts to answer the first of these two questions have been neglected in studies of instrumental social learning, and hindered by two issues: the conflation of imitation with high-fidelity cultural transmission, and the conflation of imitation with emulative learning.

Imitation is oft assumed synonymous with the successful or high-fidelity transmission of instrumental culture, whilst other mechanisms, such as local enhancement or end-product emulation, are assumed to result in lower-fidelity, less successful learning outcomes (e.g., Caldwell et al., 2012; Tennie et al., 2009). However, there is little empirical evidence that supports the relationship inferred, which may have resulted in part from an oversight of imitation's dual cultural role.

Specifically, in one of its roles, imitation might act to *facilitate* the transmission of instrumental culture (e.g., creating technological artefacts, or food extraction). Alternatively, imitation might *comprise* the cultural behaviour (as in the case of imitating many ritual and conventional behaviours; see below). In its latter role, successful or high-fidelity cultural transmission is synonymous with the occurrence of imitation. Yet in its former role, high-fidelity cultural transmission does not automatically infer imitation, or vice versa; though this is frequently assumed.

Experiments utilising the two-action task exemplify the conflation between imitation and high-fidelity learning. To my knowledge, only one study utilising the two-action paradigm has specifically attempted to address whether behaviour indicative of imitation is associated with higher levels of successful learning, relative to lower-level learning mechanisms. Interestingly, Fredman and Whiten (2008) report that, during a tool learning task, groups of capuchin monkeys displaying a high level of matching to the demonstrated actions were not significantly more successful at obtaining a food reward than monkeys limited to learning through stimulus enhancement. This finding chimes with examples of behavioural transmission in more naturalistic settings. Both the spread of potato washing behaviour in Japanese macaques, and of milk-bottle-top opening through populations of British blue tits, were initially reported as 'cultural' behaviours transmitted by processes akin to human imitation (Kawai, 1965; Bonner, 1980; Manning, 1979; Wittenberger, 1981), yet later re-described to more plausibly arise from simpler learning mechanisms, such as stimulus or local enhancement (e.g. Galef, 1990, 1992; Heyes, 1994; Laland & Hoppitt, 2003; Tomasello, 1999). Whether or not, and in what contexts, imitative learning leads to an increase in the successful transmission of instrumental culture remains an open empirical question.

The two-action task, regarded as the 'litmus test' of imitative learning (Caldwell & Whiten, 2002), also demonstrates a further conflation in debates about human cultural learning: that of imitative versus emulative learning. Evidence of matching the demonstrated solution (i.e., of moving the apparatus, or of using a particular body part) does not discriminate between whether an observer was matching the demonstrator's body actions (imitation), object movements (emulation), or both (Heyes & Ray, 2000). The difficulty inherent in teasing these two processes apart has led some to call into question the utility of conceiving an imitation-emulation dualism in instrumental learning, which could represent an oversimplified dichotomy of the action- and object-oriented processes required (e.g., Heyes, 2012a; Whiten et al., 2004). Proponents of the imitation-emulation dichotomy, who emphasise a special role for 'imitative' learning in humans (e.g., Tennie et al., 2009), have also fallen into this trap, pitching the human-nonhuman cultural divide in terms of 'process copying' (humans) versus 'product copying' (nonhumans). Specifically, it is unclear whether copying the process of action movements (imitation), or the process of object movements (emulation), is of greater importance in this context.

Interestingly, in this vein, experiments employing 'ghost conditions' specifically to isolate and test emulative learning abilities suggest that human children are more competent than chimpanzees at emulating object movements, despite the latter's characterisation as 'emulators' (Hopper et al., 2007, 2010). In contrast, when provided with information about the full process (action and object movements), matching to a demonstrator's method of task manipulation has been evidenced across a range of primate species, from chimpanzees (Whiten et al., 2005), to capuchins (Dindo et al., 2010), colobus monkeys (Price & Caldwell, 2007), marmosets (Voelkl & Huber, 2007), and lemurs (Stoinski et al., 2011), suggesting that faithfully

matching the demonstrated process is not an exclusive human proclivity (see also Whiten et al., 2009). Yet as noted above, whether or not this ability results in higher fidelity transmission of learning outcomes (i.e., instrumental success) has been largely overlooked, but is considered in more detail, using data from published two-action tests of primate social learning, in **Chapter 2**.

Recent empirical attempts to examine whether imitation is required for cultural ratcheting have relied on laboratory simulations of cumulative learning, within transmission chains of human participants tasked with instrumental goals (e.g., building a paper airplane). These studies have also conflated imitative versus emulative learning processes, comparing the efficacy of opportunities to copy artefacts (end product emulation) with opportunities to copy the construction process (process imitation/emulation). The results of these experiments are equivocal, however, with some studies claiming that learning from the end products of others is sufficient to generate cumulative improvements (Caldwell & Millen, 2009; Zwirner & Thornton, 2015); another study claiming that process information is required (Wasielewski, 2014); and yet another study suggesting that opportunities to witness the construction process might not be sufficient to sustain cumulative improvements, unless complemented by verbal instruction (Morgan et al., 2015a). The inconsistency of findings across studies likely results from a combination of procedural differences (e.g., whether or not participants received explicit information about the efficacy of previous task solutions) and differences in the novelty and difficulty (opacity) of the tasks presented (Caldwell & Millen, 2010; Derex et al., 2012), implying that the utility of opportunities to copy action and/or object movements in cumulative learning is likely context specific.

Only one transmission chain study has attempted to examine the effect of learning mechanisms on the fidelity with which artefact designs were transmitted between successive chain generations. Zwirner and Thornton (2015), employing a task that required participants to make a basket, found that learning from end products resulted in lower transmission fidelity (of basket design) relative to opportunities to witness others building or to transmit information through verbal instruction, but did not prevent cumulative improvements in basket success (i.e, ability to carry rice). It is unclear why learning from end products resulted in lower fidelity transmission of basket design in this study, but the findings do highlight the interesting possibility that some cultural artefacts, through innovative redesign, might not depend upon the highest level of transmission between learners in order to gain gradual improvements in efficiency or function. However, Zwirner and Thornton's findings also supported an especial role for high-fidelity teaching through verbal instruction – chains that permitted teaching trumped all other learning mechanisms with regards basket efficiency – echoing results from another transmission chain study that investigated the real-world and difficult-to-acquire skill of stone tool knapping (Morgan et al., 2015a).

Teaching, through verbal instruction, is one of a suite of socio-cognitive traits, including shared intentionality and extreme pro-sociality, that are considered unique or much enhanced in humans, and which, over and above our propensity to engage in imitative learning (and/or emulative learning), are believed to play a key role in cultural ratcheting (Tomasello, 1999). While it has been suggested that intention attribution might also be involved in guiding nonhuman social learning (Burkart et al., 2012), the absence of evidence for intentional teaching in other animals (Tomasello & Herrmann, 2010), coupled with evidence suggestive of their reduced proclivity to

engage in cooperative endeavours relative to our own (e.g., Warneken et al., 2006), lends some support to this view. So too does the observation that human children appear particularly receptive to pedagogical cues (Gergely & Csibra, 2006), readily incorporating the intentional teachings of adults, and actively perpetuating the trans-generational fidelity of their culture (Tomasello & Herrmann, 2010). Perhaps cementing the evidence that a suite of cultural cognition, rather than any one single cognitive trait, is what equips humans for cumulative learning (Tomasello, 1999), is a study directly comparing the learning abilities of children, chimpanzees and capuchins. Using a sequential puzzle box, exhibiting three stages of increasing difficulty, Dean et al. (2012) showed that only the children were able to build upon their achievements at lower levels. Critically, the key to their success appeared to be enhanced tendencies towards pro-sociality, verbal teaching, and imitative/emulative learning, relative to the chimpanzees and capuchins.

Imitation in ritual and conventional behaviour

Complex and cumulative culture also often depends upon conventional and ritual behaviours, and it is in these that imitation occupies an indisputable role. As in the case of transmitting gestures, dance and many other arbitrary cultural expressions, imitation itself becomes the end goal of cultural learning, and the fidelity with which action topographies are transmitted is paramount. This observation has led many authors to argue that high-fidelity motor pattern imitation might have evolved to function more predominantly as a 'social glue' rather than in instrumental cultural propagation (Clegg & Legare, 2015; Heyes, 2012a). This suggestion finds support in evidence from nonhuman animals, who appear to imitate action topographies more accurately when tasked with recreating arbitrary body movements than when

attempting to manipulate objects to recreate instrumental outcomes, but also experience increased difficultly when tasked with the former compared to the latter goal (reviewed in Huber et al., 2009).

Additional support for this view is offered by a series of recent studies with human children, examining the fidelity with which they imitate a sequence of object manipulations. These studies have demonstrated that children's imitative fidelity increases when the demonstration is provided using a ritual rather than instrumental stance (i.e., synchronous vs. non-synchronous behaviour; Herrmann et al., 2013), or using conventional rather than instrumental language cues (Clegg & Legare, 2015; Legare et al., 2015; Herrmann et al., 2013), and also when they have been ostracized rather than included by members of their in-group (Watson-Jones et al., 2015). Legare and Nielsen (2015) argue that this reduction in imitative fidelity that occurs when learning is instrumentally rather than conventionally driven is necessary for the discovery and cultural accumulation of novel technological inventions.

Over-imitation

'Over-imitation' describes the fascinating phenomenon whereby both children and adults, in experimental settings, have been observed to blanket copy even those parts of an action sequence that are manifestly causally irrelevant to achieving an instrumental goal (e.g., tapping a tool on a puzzle box, or operating a task manipulandum that is visibly disconnected from the instrumental goal; Lyons et al., 2007; McGuigan et al., 2011). This tendency, which on the surface appears to demonstrate an 'irrational' and potentially 'costly' copying mechanism (Price et al., 2010; McGuigan et al., 2007), has been observed across diverse cultures (e.g. Nielsen

& Tomaselli, 2010; Nielsen et al., 2015), but has not been replicated in our closest extant relative, the chimpanzee (Horner & Whiten, 2005).

Debates concerning this puzzling phenomenon exemplify imitation's dual role. Some authors have offered explanations grounded in causal cognition, arguing that over-imitation may reflect an adaptive high-fidelity copying mechanism, which facilitates our species unique capacity for cultural ratcheting through ensuring the transmission of causally opaque artefact functions; but which also occasionally misfires resulting in non-efficacious behaviours being copied (e.g., Horner & Whiten, 2005; Lyons et al., 2007). Others have argued that, while the experimental paradigms used to study over-imitation reflect an historical emphasis within comparative psychology on the instrumental functions of imitation, such high-fidelity copying is more plausibly explained by socially driven motivations, such as affiliative or conventional goals (e.g., Legare et al., 2015; Nielsen, 2008; Over & Carpenter, 2013; for a detailed overview of the competing hypotheses see Kenward, 2012). The use of instrumental puzzle boxes in over-imitation research has again resulted in the conflation of imitative learning with other mechanisms such as emulation, leading some authors to propose 'over-copying' as a more appropriate term (Whiten et al., 2009).

There has been little consensus reached in the ongoing debate. At one extreme, authors have argued that irrelevant actions are automatically encoded as causally relevant, and that their performance is not responsive to behavioural cost (Lyons et al., 2007, 2011). At the other extreme, it has been claimed that irrelevant action copying is intentional, and that it is moderated by social cues such as normativity (e.g., Keupp et al., 2013, 2015) and by the perceived behavioural costs of performing the action (Keupp et al., 2016). Experiments are required to clarify a number of

important questions. These include: How do selective and adaptive learning biases (see below) impact the transmission of causally irrelevant actions? Is 'over-imitation' prevalent in social learning in the 'real world', or is it predominantly an artefact of experimental design? And, To what extent does the term 'over-imitation' – which might suggest a natural predilection to 'too much copying' – accurately convey the copying propensities of our species? I attempt to address and discuss some of these questions in **Chapter 3**.

1.4. Strategic and biased social learning

The puzzling phenomenon of 'over-imitation' draws our attention to the possibility that social transmission might not always be beneficial: information might be out-dated, misleading or ineffective (Giraldeau et al., 2002; Rogers, 1988). When individual trial-and-error learning is costly or risky, high-fidelity social transmission offers an alternate, potentially cheap mode of valuable information acquisition: but only if learning is directed towards reliable social models (Boyd & Richerson, 1985; Rendell et al., 2010b; Kendal et al., 2009a). In order to solve this conundrum, researchers interested in cultural evolutionary processes have focussed on how animals filter social information adaptively, such that they might take advantage of the information available from others. The solution proposes that natural selection has fashioned social learning that is strategic, and guided by adaptive heuristics regarding *'when'* and *'whom'* and '*what'* to copy (Boyd & Richerson, 1985; Henrich & McElreath, 2003; Laland, 2004).

'*When*' strategies refer to the conditions under which an individual might choose to learn socially instead of individually, and include heuristics such as, '*copy when uncertain*,' or, '*copy when individual learning has proven too costly, or dissatisfying, or failed*' (e,g., Laland, 2004). Evidence supporting the existence of '*when'* strategies has been provided by theoretical models of cultural evolution, which suggest that social learning may only be preferred when individual learning has proven unviable (e.g., Boyd & Richerson, 1995; Enquist et al., 2007; though see Rendell et al., 2010a), and by empirical investigations of social learning, conducted with both human (e.g., Morgan et al., 2012; Toelch et al., 2014) and nonhuman animals (van Bergen et al., 2004; Webster & Laland, 2008).

'*Who*' and '*what'* strategies, also referred to as learning biases (Boyd & Richerson, 1985), describe whom one should choose to copy or learn from and what one should copy. These have also received support from a number of theoretical models (e.g., Henrich & Boyd, 1998; Powell et al., 2009), and from empirical studies conducted with humans and non-human animals (e.g., Kendal et al., 2015; Morgan et al., 2012). '*What*' strategies can be based on the content of the information conveyed, such as content that evokes a strong emotional response (e.g, Bell & Sternberg, 2001) or conveys useful social or survival information (e.g, Mesoudi et al., 2006). '*Whom*' strategies depend on the learning context, such as properties of the model's age (e.g., Wood et al., 2012), success (e.g., Mesoudi, 2008), and familiarity (e.g., Corriveau & Harris, 2009), or on the relative frequency of models (e.g., Morgan et al., 2015b), or frequency of the behaviour (e.g., Toelch et al., 2010), in the population.

The existence of guided or biased social learning not only serves to steer individual social learning away from out-dated or inappropriate information, but also profoundly influences the process of cultural evolution at the population level, across

spatial and temporal dimensions. For example, learning biases such as 'copy the most successful' or 'copy the majority' serve to increase group-level behavioural coordination, through directing cultural evolution towards homogenisation (Chudek & Henrich, 2011; Kendal et al., 2009a), and are also believed to occupy a key role in human cultural ratcheting, through guiding population-level learning towards the currently established local optima (Boyd & Richerson, 1985; Henrich & Boyd, 1998; Powell et al., 2009; for a discussion of how these biases might also act to reduce the tempo of cultural ratcheting see Eriksson et al., 2007; Mason et al., 2008)

Particular attention, from theoreticians and empiricists alike, has been awarded to majority-biased transmission (van Leeuwen et al., 2015), which is examined in relation to 'over-imitation' in **Chapter 3**. Because the majority behaviour represents the behaviour that the greatest proportion of group members have converged upon, it is expected to signal a relatively safe, reliable, and adaptive behavioural response (Boyd & Richerson, 1985; Henrich & Boyd, 1998; Hastie & Kameda, 2005; Wolf et al., 2013). For a majority trait to remain stable, the probability with which a naïve individual adopts the trait must at least match the proportional size of the majority group within the population. When the probability that the majority trait is adopted is disproportionate, or exceeds the proportional size of the majority, it is referred to as conformist transmission, and the trait is expected to increase in the population to fixation (Boyd & Richerson, 1985).

Theoretical investigations suggest that even when transmission fidelity is very high, the error introduced by just a very small number of copying imperfections can result in a substantial decay in the longevity of cultural traits, unless augmented by stabilising biases, such as majority or conformist transmission (Claidière & Sperber, 2010). The absence of majority-biased copying tendencies has been proposed to

explain the often fleeting and temporary establishment of arbitrary behavioural traditions in nonhuman animal societies, and is illustrated by a field experiment with wild meerkats (Thornton & Malapert, 2009). However, two recent field experiments, investigating artificially induced feeding preferences in wild vervet monkeys (van de Waal et al., 2013) and birds (Aplin et al., 2015b), and an agent-based simulation of natural communication patterns in sperm whale clans (Cantor et al., 2015), have challenged the view that nonhuman animal cultures lack the conformist-like tendencies necessary to establish and maintain behavioural traditions; echoing earlier experimental findings from captive groups of chimpanzees (Whiten et al., 2005), and fish (Pike & Laland, 2010). (For a discussion of whether the behavioural patterns observed are strictly attributable to majority-biased or conformist learning strategies, see Aplin et al., 2015a; van Leeuwen et al., 2015.)

Although social learning strategies are typically deployed in a mechanism neutral manner (Laland, 2004), a recent series of articles (e.g., Heyes & Pearce, 2015; Heyes, 2016) has nonetheless taken issue with the depiction of social learning strategies as cognitively fixed or hardwired algorithms. The work presented in **Chapter 3** suggests that, on the contrary, many social learning strategies are likely highly flexible, and that their operation could be influenced by a range of developmental and context specific interactions, including cultural processes in humans (reviewed in Mesoudi et al., 2016). Future work on social learning strategies must look to identify how heuristics governing the use of social information differ between human and nonhuman animals in ways that help explain the cultural gap. Simulations already undertaken on this issue have suggested that increases in highly selective and refined copying strategies, along with a greater reliance on social information more generally, are two such candidate factors (Rendell et al., 2010a).

1.5. Large-scale human interaction, cooperation and culture

The study of biased or strategic social learning illustrates the importance of contextualising the operation of cultural cognition at the level of cultural groups. Exemplifying this point further is the perplexing disconnect between the emergence of modern human culture and modern human brains. While there is much debate concerning the exact spatial and temporal emergence of cultural modernity (Klein, 2000; Mcbrearty & Brooks, 2000; Mellars, 2005; Stringer, 2007) – which includes the first evidence of consistent symbolic behaviour, sophisticated hunting technology, and artefacts constructed from materials such as ivory, bone and antler – its heterogeneous spatial and temporal structuring, frequently reported as occurring many tens of thousands of years after human brain size reached anatomical modernity ~160-200 kya (Stringer, 2007), raises an unavoidable question: How do we account for the misfit between the chronological emergence of modern human culture and modern human brains?

Building on earlier work by Henrich (2004a), Powell et al. (2009), combining information from the archaeological record, population genetics and agent-based simulations of cumulative cultural transmission, have demonstrated that demographic processes, such as increased population density and increased migration between sub populations, were likely a major determinant in the onset of cultural modernity, and can also account for the heterogeneity of its emergence across space and time: Only after having reached a critical population density threshold are cultural groups able to maintain sufficient cultural fidelity to initiate the ratcheting process. Powell and colleagues' findings are intuitive: larger numbers of highly skilled demonstrators are likely to exist in larger populations, buffering groups against stochastic population events and information loss through erroneous transmission, if individuals

preferentially attempt to learn from highly skilled demonstrators. Empirical support for these findings has been offered both from 'real world' analyses of cultural complexity among island populations in Oceania (Kline & Boyd, 2010), and from computer-based laboratory experiments of cumulative learning (Derex et al., 2013; Derex & Boyd, 2016; for a contrary view see Vaesen et al., 2016).

That humans are able to exist in such large-scale societies and engage in the frequent exchange of cultural information between non-relatives hinges upon yet another unique tendency: our ultra-cooperativeness. Indeed, observations of human cooperativeness have long presented an enigma for evolutionary biologists. The accompanying literature is vast, and driven by debate and disagreement surrounding whether standard gene-based evolutionary theories – such as direct reciprocity, indirect reciprocity and kin selection – can account for the seemingly extreme levels of cooperativeness found within large and primarily unrelated groups of humans (Fehr et al., 2002; Boyd & Richerson, 1985; Gintis et al., 2003; West et al., 2011). Semantic inconsistencies surrounding the use of terms such as *altruism*, *mutualism and strong reciprocity* in the study of cooperation have further compounded the debate over its evolution (West et al., 2007). This debate is beyond the scope of this thesis and therefore will not be debated here (for recent reviews see: Richerson et al., 2016; West et al., 2011). However, in discussing cooperative behaviour, I adopt a simple, inclusive working definition, which is popular in the empirical literature (West et al., 2007), that includes: 'any behaviour that provides a benefit to the recipient, which could be beneficial or costly to the actor,' Sachs et al., 2004).

A related topic that is of much relevance to this thesis is the growing evidence of a mutual dependence between cooperation and culture. For instance, a network-based simulation of cultural accumulation across extant ape and early human

populations, parameterised by estimates of each species' life history and socio-ecology, suggests that increases in cooperative interactions, which were possibly initiated by increases in cooperative breeding, were a driving force in the initiation of cultural ratcheting, through enhanced opportunities for social learning and exchange (Pradhan et al., 2012). This echoes findings from comparative analyses across extant primate species, showing that cooperative breeding is positively associated with proactive prosociality, social transmission and more complex communication; traits that are considered particularly enhanced in the suite of human cultural cognition (Burkart & van Schaik, 2010; Burkart et al., 2014).

Even more compelling support for the coevolution of cultural complexity and cooperation is provided by investigations of human interaction in the 'real world'. Hill et al. (2011), in an examination of human social networks across 32 contemporary foraging societies, discovered the signature of a uniquely human pattern of social interaction. Contrary to previous beliefs, the hunter-gatherer bands that constitute these 32 diverse societies exhibit high levels of co-residence between non-kin – a pattern not found in other primates – which facilitates enhanced cooperation and cultural exchange within and between neighbouring bands. Hill and colleagues argue that the large interaction networks between non-kin that ensue, drive increased information sharing between bands, amplifying the evolutionary processes underpinning human cultural ratcheting, and engaging culture and cooperation in a coevolutionary cycle.

A subsequent analysis of Ache and Hadza hunter gatherer tribes reinforces these findings, showing a high probability of cultural and cooperative interactions between any two randomly chosen members of a tribe, and estimating that men are likely to observe over 300 other men making tools throughout their lifetime, in

comparison to male chimpanzees who are only likely to interact with 20 other males (Hill et al., 2014). Moreover, a social network analysis of Tanzanian hunter-gatherers has revealed striking structural regularities between the network structures of hunter-gatherers and modernised populations, raising the possibility that natural selection has fashioned some structural properties of human social networks (Apicella et al., 2012). The consistent network properties this analysis identified – transitivity and reciprocity between individuals – are believed to foster connections between non-kin and ensure that co-operators tend to cluster together; features believed to favour the evolution of large-scale cooperation (Nowak et al., 2010) and amplify the process of cultural transmission (Hill et al., 2011).

While large-scale cooperation acts to scaffold the evolution of increasing cultural complexity, cultural transmission functions to stabilise and maintain cooperative interactions. Culture's role in maintaining societal cooperation can be seen most conspicuously in the operation of socially transmitted norms of cooperative behaviour, and in the cultural evolution of formal institutions, such as central governments and religious laws (Chudek & Henrich, 2011). Indeed, young children readily learn and incorporate, and are also strikingly inflexible about, the normative behaviours and values of their culture (Levy et al., 1995; Tomasello & Herrmann, 2010). Moreover, a series of experiments conducted using economic games suggest that norms of cooperation vary cross-culturally (Henrich et al., 2005; Herrmann et al., 2008). These experiments, that have included ultimatum, third-party punishment, and public goods games, indicate that individuals across all societies cooperate at higher levels than if they were trying to maximise their own economic gain, but also vary in cooperativeness in a manner consistent with the institutional and civic properties of their cultures. This variation has fuelled wide debate regarding whether cultural group

selection – which occurs if groups that evolve highly cooperative norms are able to outcompete those that didn't – has played a key role in the evolution of human hyper-cooperation (reviewed in Richerson et al., 2016).

At the opposite end of the spectrum, the mechanisms of simple and automatic imitation (see '*Imitation*' above) are also believed to be hard at work in cementing cooperation within cultural groups. Humans have long been observed to subconsciously imitate each other's body movements and mannerisms; a tendency that has been replicated in semi-natural experimental settings, where it has also been shown that its occurrence serves to promote cooperative interactions (reviewed in Chartrand & Van Baaren, 2009). Automatic imitation can even occur when it interferes with our ability to perform a task (e.g., our ability to grasp an object when watching another simultaneously releasing their grasp; Heyes, 2011); a phenomenon that so far has only been replicated in budgerigars and domestic dogs (Range et al., 2011). It is proposed that, just as our predilection towards subconscious imitation fosters increased liking and cooperation, cooperative interaction in turn serves to increase imitation; thus engaging these two processes in a virtuous circle of reciprocity (Heyes, 2012a). Although the mechanisms that have given rise to this virtuous circle in human interactions are still poorly understood, it is plausible that the processes responsible for engaging simple imitation and cooperation in this process of mutual reinforcement, have not only acted as a 'social glue', but also as a springboard for more complex forms of cultural learning and cooperation (Heyes, 2012a). The virtuous circle is discussed further and forms the basis of the work presented in **Chapter 4.**

Between the formal dictations of our cultural norms and institutions, and the silent operation of the virtuous circle, there is one other obvious, yet overlooked, way

that our cultural dependence might solidify cooperation: reputation. A fundamental property of human interaction involves repetition: and where repetition is possible, reputation is often at stake. The public goods game (PGG), which exemplifies the interplay between cooperative behaviour and reputation, is the classic laboratory paradigm for studying the collective action dilemmas that routinely characterise human interactions. In a typical PGG experiment, individuals choose how much to contribute to a public pool, which benefits all group members equally. The maximum benefit for the group is then achieved if all group members contribute the full amount; yet individual free riders benefit more by withholding their own contribution while still benefiting from the pool.

Despite the economic incentive to free ride, behaviour in the PGG has been used to argue that humans have evolved strong prosocial tendencies (Fehr et al., 2002; Rand et al., 2012; although see Burton-Chellew & West, 2013). For instance, cross-culturally human subjects contribute at much higher levels to the PGG than if they were maximising their individual gain (Henrich et al. 2005). They are also willing to punish defectors (Henrich et al., 2006), and reward co-operators (Rand et al., 2009), at an additional cost to themselves, which serves to discourage the emergence of free rider strategies in repeated interactions. Cultural norms appear to play a role, and reputation is key. Yet cultural exchange might often constitute a cooperative dilemma in its own right (Cabrera & Cabrera, 2002; Rendell et al., 2011b) – of sharing individually-gleaned information for the collective benefit, or withholding that same information to maintain a competitive edge – further helping to consolidate the reputation-based matrix that upholds large-scale interaction. Moreover, if access to cultural information is itself dependent on the establishment of cooperative reputation, then it is plausible that our species' extreme dependence on cultural learning may

serve to encourage cooperative reputation building across multiple behavioural domains; potentially offering a cheaper system of sanction or reward against free riders than other costly alternatives. This suggestion will be examined further in **Chapter 5**.

1.6. Summary

The glaring and unparalleled complexity that sets human culture apart from the 'cultures' of other animals – our artefacts, rituals, social institutions and technologies – is underpinned by a complex of interacting processes, that we are only just beginning to understand. At the micro level, the specific psychological and behavioural mechanisms that afford social learning, especially the faithful transmission of instrumental culture, are still a moot point. Yet faithful copying is ineffective unless directed towards the right opportunities and individuals. A growing body of evidence is beginning to unravel the learning heuristics that guide faithful copying; that permit us to filter and take advantage of social information, and preserve fundamental facets of our cultural repertoire. The operation of learning heuristics reminds us that faithful and guided transmission plays out in populations, both within and between cultural groups. Cultural evolution on the macro level is dependent on our remarkable ability to cooperate in large groups with non-kin. The evolution of our ultra-cooperativeness, in turn, has likely been scaffolded by its interaction with the mechanisms and processes underpinning cultural transmission, yet this relationship is currently under-researched. In sum, a deeper understanding of culture's constituent mechanisms, how they interact, support, amplify, and constrain the cultural process,

from the micro through to the macro level, is key to understanding the unparalleled feats of human culture.

In the following chapters, I first attempt to contribute to outstanding questions concerning the role of social learning mechanisms, and of social learning heuristics, in the evolution of high-fidelity cultural transmission. The focus of the thesis then shifts towards an examination of the interaction between social transmission and cooperation, and of the role this interaction might play in supporting cultural complexity in humans. Specifically, in **Chapter 2,** I present the results of a meta-analysis undertaken on studies of primate social learning abilities that were conducted using the two-action paradigm. In doing so, I attempt to shed light on the [hitherto neglected] question of *if* and *when* fidelity to the demonstrator's behaviour (i.e., imitation or emulation) is required for the faithful transmission of instrumental outcomes. **Chapter 3** focuses on 'over-imitation' in young children. Taking a cultural evolutionary approach, I examine the operation of adaptive learning biases (here of majority-biased copying) in response to behaviour containing causally irrelevant actions. I explore whether children's copying of the majority extends to majorities that perform irrelevant actions, and discuss whether the depiction of humans as 'over-imitators' is appropriate, in light of my findings.

In **Chapter 4**, I present a series of experiments that explore the hypothesised operation of a 'virtuous circle' between simple imitation and cooperative behaviour. In this chapter, I focus on whether the establishment of cooperative rapport between interaction partners subsequently serves to increase imitative behaviour – a relationship that has previously received little empirical investigation. **Chapter 5** explores the interaction between human cultural learning and the mechanisms sustaining large-scale cooperation, using an experimental micro-society approach, and

the adoption of two linked computer games. Specifically, in this chapter I ask whether individuals' dependence on access to cultural information can act to stabilize cooperative reputations and behaviour during repeated interactions in a public goods dilemma. Finally, in **Chapter 6**, I end with a general discussion of the findings, and of some wider implications for the study of social learning and cultural evolution.

Chapter 2

High-Fidelity Social Transmission: A Meta-Analysis of Social Learning Studies Across Primate Species

Abstract

Although it is commonly assumed that learning mechanisms associated with increased fidelity to a demonstrator's actions and/or object manipulations (i.e., imitation and/or emulation) are required for the successful transmission of instrumental cultural goals, there is currently little evidence to support this claim. In this chapter I present the findings of a meta-analysis that was undertaken on published studies of primate social learning abilities that were conducted using the two-action paradigm. The results of this analysis suggest that, relative to learning through other mechanisms (e.g., local/stimulus enhancement, individual learning), social learning that exhibits high fidelity to the demonstrator's behaviour offers learning advantages that increase incrementally with the complexity of the learning goal. This finding supports claims that learning mechanisms generating high demonstrator fidelity play a key role in the emergence, maintenance and cumulative improvement of culture.

2.1. Introduction

The fidelity of social transmission – the accuracy with which learned information and behaviours spread between individuals – is considered pivotal to the extraordinary and cumulative cultural success of our species (e.g., Boyd & Richerson, 1985; Galef, 1992; Tomasello et al., 1993; Lewis & Laland, 2012). It has been argued compellingly that the faithful and efficient transfer of culturally learned information is what allows each new generation of human cultural learners to build upon the achievements of its predecessors – a process termed 'cultural ratcheting' – and that this ability is lacking or much reduced in other species (Tomasello, 1999). This proposal has received support from theoretical models of cultural evolution (Enquist et al., 2010; Lewis & Laland, 2012), and has fuelled debate among interested scholars over whether other animals, particularly nonhuman primates, share the imitative learning capabilities frequently attributed to humans (Whiten et al., 2004).

While theoreticians have focussed on modelling fidelity explicitly as a function of learning outcomes (i.e., the frequency, complexity, perseverance or payoffs associated with cultural traits; e.g., Enquist et al., 2010; Rendell et al., 2010a; Lewis and Laland, 2012), experimental psychologists have focussed on defining and detecting the psychological mechanisms and behavioural processes believed to underpin the fidelity of cultural transmission. A suite of socio- or cultural-cognitive abilities have been proposed – including verbal teaching, extreme prosociality, and imitation – that are considered unique or much enhanced in humans (Tomasello, 1999), and which have received some empirical support for their concerted role in human cultural ratcheting from a recent comparative study of sequential learning in chimpanzees, capuchins and children (Dean et al., 2012).

However, the majority of investigations of social learning abilities, and of attempts to elucidate the uniqueness of human social learning, have focussed on experimentally testing for imitation. In contrast to teaching, which has proven especially difficult to verify in other species, despite occupying a seemingly diverse taxonomic distribution (Hoppitt et al., 2008), and human language, which is regarded as qualitatively different from the communication systems of other species (e.g., Hurford, 2007; Scott-Phillips, 2014), the ability to engage in imitative learning has to a lesser degree been considered the *fait accompli* of human beings (e.g., Voelkl & Huber, 2000; Whiten et al., 2004, 2009). Indeed, following an upturn in discussions and debates concerning whether human social learning might be uniquely imitative among the great apes (Galef, 1990; Tomasello, 1990; Heyes, 1993; Whiten et al., 2004, 2009), there has been a flurry of systematic investigations of imitative learning abilities in human and non-human primates during the last two- to three-decades.

As detailed in the General Introduction (*Section 1.2.*), most researchers have adopted the two-action task to test for imitative or 'high-fidelity' learning abilities (Dawson & Foss, 1965). Subjects are offered two alternative solutions for solving a puzzle (i.e., retrieving a reward), but half witness the demonstration of one solution (e.g., push the lever to the left), while the other half witness the alternate solution (e.g., push the lever to the right). Imitation or 'high-fidelity' learning is then inferred if the test group demonstrate a disproportionate tendency to perform (i.e., match) the solution they observed with a probability greater than expected at chance.

Despite being hailed as the litmus test of imitative learning, the two-action paradigm does not discriminate between the relative importance of copying the action topography of the demonstrator's body movements (imitation), and reproducing the object movements that resulted from the demonstrator's actions (emulation) (Heyes &

Ray, 2000). Yet while this discrimination has been deemed important by authors who emphasise an especial role for action imitation in human cultural affairs (e.g. Tomasello and Call, 1997), other authors have questioned the utility of awarding action coping an elevated status in instrumental cultural learning, which plausibly benefits from both action- and object-oriented copying processes (Whiten et al., 2004, 2009). Still other authors have suggested that copying object movements (i.e., emulation) might be more important in transmitting instrumental culture than copying actions (Heyes, 2012a), and that detailed action imitation may instead be the preserve of social rather than instrumental goals (Legare & Nielsen, 2015).

While the two-action task does not speak directly to the ongoing debate over learning mechanisms, if positive results generated using this paradigm (i.e., evidence of matching the demonstrated solution) do provide an indication of high-fidelity learning abilities, then it becomes apparent that this ability is not exclusive to humans: a range of taxonomically diverse species are also implicated, including many species of primate (e.g., Dindo et al., 2008; Stoinski et al., 2011), dogs (Range et al., 2007), and birds (Akins & Zentall, 1996). Yet while researchers have been busy applying the two-action paradigm across a range of taxa, they have largely neglected to confirm that evidence of matching the demonstrated behaviour is also concurrent with measurable increases in the faithful transmission and acquisition of the learning goal or outcome (e.g., reward retrieval or the construction of material culture).

Only one study has directly examined this relationship, but found no evidence that fidelity to the demonstrator's behaviour increased transmission of the learning outcome (reward retrieval), relative to lower-level learning mechanisms, such as stimulus enhancement (Fredman & Whiten, 2008). This finding chimes with observations of the successful spread of potato washing behaviour in Japanese

macaques, and of milk-bottle-top opening through populations of British blue tits, which despite being initially attributed to processes akin to imitation, later emerged to more plausibly arise from simpler learning mechanisms, such as enhancement effects (e.g. Galef, 1990, 1992; Laland & Hoppitt, 2003). Transmission chain studies conducted with human participants, which have attempted to assess the role of learning mechanisms in the cumulative improvement of laboratory artefacts, have also returned inconsistent conclusions. For instance, opportunities to observe the construction behaviour of more experienced individuals (i.e., engage in imitative and/or emulative learning), relative to opportunities to learn from other's end products, were not required to produce cumulative improvements in two studies (Caldwell & Millen, 2009; Zwirner & Thornton, 2015), but were deemed necessary in a third (Wasielewski, 2014).

There is a clear need to assess further *if* and *when* fidelity to a demonstrator's behavioural process (i.e., *demonstrator fidelity*) is required for the successful and faithful cultural transmission of instrumental and material outcomes (i.e., *outcome fidelity*). In the current study, I attempted to address this question by conducting a meta-analysis on the growing body of primate social learning experiments that have utilised the two-action task paradigm. Despite researchers having largely overlooked the relationship between learning mechanisms and learning outcomes, it is convention to report both measures in published manuscripts, which presented the opportunity for my analysis.

Some authors have extended the two-action task to incorporate multiple manipulandi in task designs of increasing complexity (e.g. Whiten, 1998; Stoinski & Whiten, 2003), and a number of task designs have also incorporated tool use as a necessary component (e.g. Call & Tomasello, 1994; Hopper et al., 2010). Tool use

among the animal kingdom appears to be exceedingly rare, reliant on simple manipulations of natural objects, and rudimentary in complexity and diversification outside the human lineage (Seed & Byrne, 2010; Shumaker et al., 2011; Biro et al., 2013). Indeed, a study conducted with chimpanzees suggests that the addition of a tool profoundly reduces performance on a task designed to test causal understanding, relative to when tool use is not required (Seed et al., 2009). Evidence from humans and non-human primate species suggests that the process of acquiring tool-using abilities is likely highly dependent upon social transmission (Biro et al., 2003; Beck et al., 2011; Stout, 2011; Price et al., 2009; Hobaiter et al., 2014), making tool use a pertinent additional component in tests of instrumental social learning.

In a preliminary analysis, using only a subset of the data (i.e. published experiments that had also included asocial learning controls), I tested whether groups that received a demonstration outperformed matched groups of asocial (individual) learners. I predicted that, across studies, social information would be advantageous, and lead to an increase in learning success relative to individual learning (Laland, 2004; Kendal et al., 2005; Rendell et al., 2010a). During the preliminary analysis, I also obtained an initial gauge of how important *demonstrator fidelity* (identified as *high* or *low* by a positive or negative result on the two-action test, respectively) was in the successful transmission of reward retrieval (*outcome fidelity*), and anticipated that while both high and low demonstrator fidelity would afford improvements in outcome fidelity over asocial learning, high-fidelity matching to the demonstrator would facilitate the greatest improvement overall (Galef, 1992; Tomasello, 1999; Dean et al., 2012).

For the main analysis, which included data from all social learning experiments meeting the inclusion criteria, I undertook a more detailed examination

of the effects of demonstrator fidelity (again identified by a positive or negative result on the two-action test) on outcome fidelity, controlling for procedural differences between experiments. I hypothesised that high-fidelity matching to the demonstrator's solution would prove more important to outcome fidelity at higher- relative to lower-levels of task complexity: a prediction that follows from the thesis that mechanisms generating demonstrator fidelity (e.g., imitation) are an essential component of the cultural ratchet (Galef, 1992; Tomasello, 1999; Dean et al., 2012; Lewis & Laland, 2012). Finally, owing to the additional learning difficulties and demands associated with tool use (Beck et al., 2011; Morgan et al., 2015b; Seed et al., 2009; Stout, 2011), I also predicted that tool-using tasks would reduce learning performance, and would benefit disproportionately from high demonstrator fidelity.

2.2. Materials and Methods

2.2.1. Search protocol

I used three search protocols to conduct literature searches in April 2014. First, keyword searches of two online databases were performed. The first 30 pages of results were obtained from Google Scholar (Google) for the search terms "*social learning and imitation in primates,"* and "*social learning and imitation in children*." Web of Science (Thomson Reuters) was also searched (in the TOPIC field) using the following keyword searches: *"imitation"* AND *"social learning"*; *"social learning"* AND *"primate*"*; *"social learning"* AND *"children*"*; *"imitation"* AND *"primate*"*; *"imitation"* AND *"children*"*; *"social learning"* AND *"tool us*"*; *"social learning"* AND *"artificial fruit"*; *"imitation"* AND *"emulation."*

Second, Web of Science was used to search for all articles citing 9 literature reviews that specifically assess and discuss the experimental evidence for imitation and other social learning mechanisms in non-human animals including primates (Galef, 1988; Whiten & Ham, 1992; Heyes, 1993; Byrne, 1995; Byrne & Tomasello, 1995; Tomasello & Call, 1997; Shettleworth, 1998; Heyes & Ray, 2000; Zentall, 2001; Whiten et al., 2004).

The titles and abstracts of all returned search results were examined. Those considered relevant were then read in detail to ascertain if they were suitable for inclusion (for study inclusion criteria, see below). Lastly, I also examined articles that were cited in the text of other articles if they had not already been identified through the online database searches.

2.2.2. Inclusion criteria

There were several criteria for study inclusion. Studies were only included if the test subjects were human children or non-human primates, restricting analyses to species within a taxonomically homogenous group. It was also essential that the study presented subjects with an instrumental learning goal (i.e., a foraging puzzle) and that success versus failure in obtaining the goal could be determined for each individual member of a test group (i.e., obtaining a food reward in the case of non-human primates or a non-food reward in the case of child participants. For two published studies (Whiten et al., 1996; Custance et al., 2006), data on the success/failure of subjects was clarified by contacting the study authors. Inclusion was further limited to articles that utilised a design compliant with the dominant two-action paradigm (outlined above), to ensure consistency in the measurement and assignment of social

learning mechanisms throughout the meta-dataset. Two studies that utilised a modified version of the two-action paradigm, to test for sequence imitation (i.e., copying of the precise order in which a series of task-directed actions has been performed; Whiten, 1998), or copying of hierarchical action structure (i.e., copying of the precise hierarchically organised series of steps of task-directed actions; Whiten et al., 2006), were also included, but were flagged due to their deviation from the predominant protocol.

Other inclusion criteria were that the test subjects should have no prior experience of the task, and that each experimental subject should be tested individually with a competent demonstrator. Two studies that did not meet the criteria of testing subjects individually were included (Horner et al., 2006; Fredman & Whiten, 2008), as their departure from this stipulation did not interfere with test subjects' opportunities to view consistent demonstrator behaviour or to access and manipulate the task.

2.2.3. Data extraction

Data were extracted from the text or tables of articles, or indirectly from figures in a small number of cases. For five articles, data were extracted from a subset of the total experimental procedures reported due to (1) procedural changes occurring in later (i.e., follow-up or second stage) test trials (Call & Tomasello, 1994; Voelkl & Huber, 2000; Whiten et al., 2006), (2) a substantial number of test subjects being unavailable to participate in later trials (Stoinski et al., 2001), and (3) the data presented being from an earlier published article (Hopper et al., 2012).

Outcome Success/Outcome Fidelity

For asocial learning control groups, *outcome success* corresponded to the proportion of individuals in the test group who successfully retrieved the reward during the experimental procedure (i.e., in any of the test trials). Likewise, for groups that received a demonstration of reward retrieval, the proportion of individuals in the group who successfully retrieved the reward during the experimental procedure represented a measure of *outcome fidelity*. For a subset of experiments, a measure of outcome success and outcome fidelity in relation to the first test trial was also obtained.

Demonstrator Fidelity

Demonstrator fidelity was categorised as 'high' or 'low' for each test group. Thus, test groups for which there was a positive result on the two-action task were categorised as 'high' demonstrator fidelity, and test groups that showed no evidence of matching the demonstrated solution were categorised as 'low'. All studies included in the meta-dataset followed the convention in inferential statistics of rejecting the null hypothesis at the 0.05 probability level (i.e., $p<0.05$), therefore this was the threshold adopted in assigning test groups to 'high' or 'low' demonstrator fidelity.

Statistical tests using the two-action paradigm are more powerful when assessing a discrimination ratio of performance of the two actions (e.g., Heyes & Dawson, 1990; Whiten et al., 1996). Accordingly, this was the statistical standard adopted when categorising test groups' demonstrator fidelity as 'high' or 'low'. However, a small subset of included experiments utilised statistical tests based on absolute rather than relative measures of action performance (Custance et al., 1999, 2001; Fredman & Whiten, 2008), or scores relating to a behavioural index of the two

actions (Price et al., 2009). For this subset, statistical evidence of intra-test-group differences in performance of both actions was required for a test group to be categorised as showing 'high' demonstrator fidelity (i.e., higher levels of action 'a' must be observed in test group subjects witnessing 'a' *AND* higher levels of action 'b' must be observed in test group subjects witnessing 'b').

Procedural variations across social learning experiments resulted in varying numbers of test trials being experienced by test groups. Hence, statistical evidence that matching the demonstrator's solution had occurred at any point during a test group's total trials was accepted, but a moderator variable relating to the number of trials test groups received was included in analyses to control for the potential influence of this factor on whether subjects learned to successfully retrieve the reward (see below). Where an apparatus contained multiple two-action manipulandi (e.g., Whiten et al., 1996; Stoinski et al., 2001), demonstrator fidelity was categorised as 'high' if the test group showed evidence of matching the demonstrated solution on at least one manipulandum.

For eight test-group-level learning assessments (six presented in Call et al. 2005; two presented in Hopper et al. 2010), I recalculated the test statistics from descriptive data reported by the authors, because statistical tests matching the criteria outlined above had not been performed. One test group (Bugnyar & Huber, 1997), in a deviation from other two-action test protocols, was compared with an asocial learning control group that demonstrated a bias towards performing the alternate solution than was witnessed by the test group. Another test group (Caldwell & Whiten, 2004) included some subjects that did not witness a competent demonstrator; thus, these subjects were excluded from the data and demonstrator fidelity in the

remaining subjects (that showed no evidence of matching the demonstrated task manipulation) was categorised as 'low'.

Apparatus designs were examined to ensure that alternate mechanisms such as stimulus or local enhancement had been ruled out when categorising demonstrator fidelity as 'high'. The fourth experiment in a battery of two-action tests conducted by Rigamonti et al. (2005) utilised a design that did not permit differentiation between 'high' and 'low' categories of demonstrator fidelity; hence all test groups included in the fourth experiment were categorised as 'low', and those that had potentially matched the demonstrator's task manipulations were flagged.

Task complexity

A measure of task complexity was devised, applicable for each apparatus, and given by the total number of causal manipulations necessary before the reward could be retrieved. For example, if test subjects were required to push or pull open a door before gaining access to a reward, this was given a task complexity rating of 1 (e.g., Tennie et al., 2006; Hopper et al., 2008). Likewise, if test subjects were required to disable two task defences in addition to removing a lid before gaining access to a reward, this was given a task complexity rating of 3 (e.g., Custance et al., 1999). Tool use during reward retrieval was coded as a separate binary variable. However, apparatus manipulations performed using a tool – e.g., using a tool to push or lift a lever (e.g., Hopper et al., 2010), or using a tool to reposition the reward inside the apparatus (e.g., Call & Tomasello, 1994) – were included in the measure of task complexity. In one experiment (Price et al., 2009), the measure of task complexity included an apparatus manipulation that was required on the tool itself (i.e., extending or sliding components of the tool together).

A separate variable detailing whether the apparatus was transparent (i.e., allowing the reward to be visible) or opaque (i.e., concealing the reward) was also extracted. Call et al., (2005) used a combination of transparent and opaque apparatuses within test groups, but found no behavioural differences relating to whether the reward was visible or not. Thus, as analyses in the current study were undertaken at the level of the experimental test group, all apparatuses in Call and colleagues' study were regarded and coded as 'transparent'.

Moderator variables

A number of additional variables relating to between-study procedural variations were extracted. The total number of learning demonstrations that subjects in each test group received, along with the total number of test trials they were awarded, were recorded. In one learning experiment (Price et al., 2009) the number of demonstrations provided to experimental subjects varied between an upper and lower limit, therefore the mid-point of this range was recorded.

A measure of each test group's mean age was also obtained. To standardise the measure of age across species, the mean age of each group was adjusted by the maximum longevity of the species, which was obtained from the PanTheria online dataset (Jones et al., 2009). In four social learning experiments (Bugnyar & Huber, 1997; Voelkl & Huber, 2000; Caldwell & Whiten, 2004; Dindo et al., 2011), test subjects were described as "adult" but information pertaining to the mean age of the test group was not given. For these four experiments, the mean age of the group was approximated, by taking the mid-point of the lifespan between sexual maturity and maximum longevity for those species, as reported in the PanTheria dataset.

2.2.4. Dataset summary

Two-action tests of learning fidelity

A total of 28 published papers detailing the results of 36 separate social learning experiments, conducted between 1993 and 2012, were included in the final dataset. Experimental subjects that did not interact with the apparatus during testing were excluded from the data, as they did not contribute behaviourally to group-level learning assessments. This resulted in 777 individual experimental subjects distributed across 67 test groups. The final dataset comprised 84 test-group-level assessments of demonstrator fidelity (nine test groups participated in more than one social learning experiment), and 926 individual-level measures of outcome fidelity (i.e., success vs. failure in achieving the instrumental goal).

Fourteen test groups were assigned to 'ghost' conditions, receiving reduced information about the actions necessary to obtain the reward (see Hopper, 2010). Four additional test groups also received incomplete information regarding the manipulations necessary to successfully operate the apparatus (Call et al., 2005), resulting in a total of 18 out of 84 test-group-level assessments of learning that followed reduced or incomplete task demonstrations.

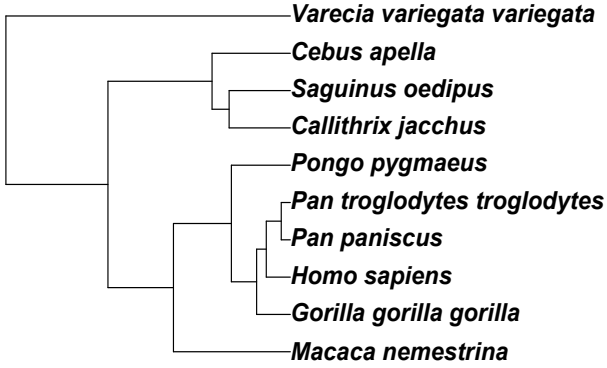
Asocial learning controls

Data on the baseline performance (i.e., success verses failure in achieving the goal) of additional asocial learning control groups were provided for only 17 of the 36 (<50%) learning experiments. A total of 249 individual control subjects distributed across 31 asocial learning groups were included in the final dataset. Control groups reported in Fredman and Whiten (2008) were excluded from the dataset, as they were provided

with some social information about the solve state of the task. Also excluded was a control group reported in Price et al. (2009) that received a partial demonstration.

Distribution of social learning tests across primate species

There were 10 different primate species listed in the dataset, including species of monkey and apes, along with Old World and New World primates. However, the majority of the data comprised social learning experiments conducted on human children (*Homo sapiens*) and chimpanzees (*Pan troglodytes*) (see Figure 2.1).



	Num of group-level learning assessments	
	Demonstration	Asocial control
Varecia variegata variegata	1	-
Cebus apella	7	1
Saguinus oedipus	2	-
Callithrix jacchus	3	3
Pongo pygmaeus	7	2
Pan troglodytes troglodytes	17	9
Pan paniscus	2	1
Homo sapiens	37	12
Gorilla gorilla gorilla	3	2
Macaca nemestrina	5	1

Figure 2.1. *The number of demonstration groups and asocial learning control groups, by species*

Apparatuses

A total of 17 different apparatuses were used across the 36 included experimental designs, indicative of some repeatability in task designs across social learning experiments but also variation. A measure of apparatus complexity was extracted (the details of which are provided above), ranging from 1 (least complex) to 7 (most complex). The majority of test-group-level learning assessments were conducted with apparatuses of the lowest complexity (see Table 2.1.).

Table 2.1. *Number of group-level learning assessments by task complexity*

	Task Complexity				
	1	**2**	**3**	**5**	**7**
Social learning test groups	61	5	12	5	1
Asocial learning control groups	21	6	1	2	1

Tool use

Eight (22%) social learning experiments involved tool use in reward retrieval. Experiments that included the use of a tool were limited to human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), capuchin (*Cebus apella*), and orangutan (*Pongo pygmaeus*) subjects (i.e., all habitual tool-using species; Ottoni & Izar, 2008; Seed & Byrne, 2010; Biro et al., 2013). Tool use was not recorded for the mother-reared test group detailed in Fredman and whiten (2008), as all subjects that successfully retrieved the reward in this test group did so without use of the demonstrated tool.

A detailed overview of the dataset following data extraction is provided in Table 2.2 (see below).

Table 2.2. *Detailed overview of the meta-dataset used in this chapter.*

Species	Group ID	PanTHERA Adjusted Age	N	N Solve	N Trials	N Demos	Full Demo	Tool Use	Task Complexity	Task Opaque	Demo Fidelity	Demo Fidelity Test statistic	Study
Pan trogoldytes	1	0.56	8	8	1	58	Yes	No	1	Yes	High	B(8/8, p=.01)	Hopper et al. 2008
Pan trogoldytes	2	0.56	8	7	1	58	No	No	1	Yes	High	B(7/7, p=.01)	Hopper et al. 2008
Pan trogoldytes	3	0.42	8	8	1	58	No	No	1	Yes	Low	B(4/7, p=.27)	Hopper et al. 2008
Homo sapien	5	0.03	8	8	1	15	Yes	No	1	Yes	High	B(7/7, p=.01)	Hopper et al. 2008
Homo sapien	6	0.03	8	8	1	15	No	No	1	Yes	High	B(8/8, p=.01)	Hopper et al. 2008
Homo sapien	7	0.03	8	8	1	15	No	No	1	Yes	Low	B (6/8, p=.15)	Hopper et al. 2008
Pan trogoldytes	9	0.21	11	10	1	1	Yes	No	1	Yes	Low	NS	Tennie et al. 2006
Pan trogoldytes	10	0.21	11	8	1	1	No	No	1	Yes	Low	NS	Tennie et al. 2006
Pan trogoldytes	12	0.14	5	4	1	6	Yes	No	1	Yes	Low	NS	Tennie et al. 2006
Pan paniscus	13	0.17	2	2	1	6	Yes	No	1	Yes	Low	NS	Tennie et al. 2006
Gorilla g. gorilla	14	0.15	2	1	1	6	Yes	No	1	Yes	Low	NS	Tennie et al. 2006
Pongo pygmaeus	15	0.14	3	3	1	6	Yes	No	1	Yes	Low	NS	Tennie et al. 2006
Pan trogoldytes	16	0.14	6	6	1	6	No	No	1	Yes	Low	NS	Tennie et al. 2006
Pan paniscus	17	0.17	1	1	1	6	No	No	1	Yes	Low	NS	Tennie et al. 2006
Gorilla g. gorilla	18	0.15	2	2	1	6	No	No	1	Yes	Low	NS	Tennie et al. 2006
Pongo pygmaeus	19	0.14	3	2	1	6	No	No	1	Yes	Low	NS	Tennie et al. 2006
Homo sapien	24	0.01	33	24	1	2	Yes	No	1	Yes	Low	B(15:9, p>.30)	Tennie et al. 2006
Homo sapien	25	0.01	24	21	1	2	Yes	No	1	Yes	High	B(18:3, p<.001)	Tennie et al. 2006
Homo sapien	26	0.02	20	19	1	2	Yes	No	1	Yes	High	B(18:1, p<.001)	Tennie et al. 2006
Homo sapien	27	0.01	29	24	1	2	No	No	1	Yes	Low	B(12:12, p>.30)	Tennie et al. 2006
Homo sapien	28	0.01	23	20	1	2	No	No	1	Yes	Low	B(12:8, p=.5)	Tennie et al. 2006
Homo sapien	29	0.02	20	17	1	2	No	No	1	Yes	High	B(16:1, p<.001)	Tennie et al. 2006

Pan trogoldytes	33	0.19	13	11	1	2	Yes	No	1	No	Low	B(2/10, p=1)	Call et al. 2005
Pan trogoldytes	34	0.19	12	9	1	3	No	No	1	No	Low	B(1/8, p=1)	Call et al. 2005
Pan trogoldytes	35	0.19	12	9	1	2	No	No	1	No	Low	B(7/11, p=.55)	Call et al. 2005
Homo sapien	37	0.02	12	11	1	2	Yes	No	1	No	Low	B(7/11, p=.27)	Call et al. 2005
Homo sapien	38	0.02	12	10	1	5	No	No	1	No	Low	B(7/11, p=.27)	Call et al. 2005
Homo sapien	39	0.02	12	5	1	2	No	No	1	No	Low	B(2/5, p=.81)	Call et al. 2005
Callithrix jacchus	41	0.47	4	0	3	3	Yes	No	3	Yes	Low	NS	Caldwell & Whiten 2004
Callithrix jacchus	44	0.47	11	11	1	75	Yes	No	1	Yes	High	T(4.5)=2.33, p=.03	Voelkl & Huber 2000
Pongo pygmaeus	48	0.33	4	1	3	3	Yes	No	5	Yes	Low	U(7,7)=5, p=.17	Custance et al. 2001
Pongo pygmaeus	49	0.07	10	0	3	3	Yes	No	5	Yes	Low	U (7,7)=5, p=.17	Custance et al. 2001
Cebus apella	50	0.26	12	12	1	40	Yes	No	1	Yes	High	U(6,6)=36, p=.005	Dindo et al. 2010
Pongo pygmaeus	52	0.28	8	8	1	10	Yes	No	1	Yes	High	$X^2(1)=6.67,p<.01)$	Dindo et al. 2011
Homo sapien	54	0.03	20	20	1	15	Yes	Yes	1	No	High	FET(p<.001)	Hopper et al. 2010
Homo sapien	56	0.03	28	24	1	15	Yes	Yes	1	No	High	FET(p<.001)	Hopper et al. 2010
Homo sapien	58	0.03	23	9	1	15	No	Yes	1	No	High	FET(p<.001)	Hopper et al. 2010
Homo sapien	60	0.03	20	12	1	15	No	Yes	1	No	Low	FET(P>.05)	Hopper et al. 2010
Pan trogoldytes	62	0.4	20	5	3	40	Yes	Yes	2	No	Low	H(4)=13.07, p=.01 H(4)=3.61, p=.46	Price et al. 2009
Gorilla g. gorilla	67	0.52	6	6	3	5	Yes	No	5	Yes	High	U(3,3)=0, p=.03 U(3,3)=0, p=.03	Stoinski et al. 2001
Macaca nemestrina	70	0.24	12	12	8	32	Yes	No	1	No	Low	B(NS)	Rigamonti et al. 2005
Macaca nemestrina	70	0.24	13	13	8	32	Yes	No	1	Yes	Low	B(NS)	Rigamonti et al. 2005
Macaca nemestrina	70	0.24	13	12	8	32	Yes	No	1	Yes	Low	B(NS)	Rigamonti et al. 2005
Homo sapien	71	0.01	10	10	8	32	Yes	No	1	No	Low	B(NS)	Rigamonti et al. 2005
Homo sapien	71	0.01	10	10	8	32	Yes	No	1	Yes	Low	B(NS)	Rigamonti et al. 2005
Homo sapien	71	0.01	10	10	8	32	Yes	No	1	Yes	High	B(Z=2.22,p=.013)	Rigamonti et al. 2005

Homo sapien	72	0.02	10	10	8	32	Yes	No	1	No	High	B(Z=2.22,p=.013)	Rigamonti et al. 2005
Homo sapien	72	0.02	10	10	8	32	Yes	No	1	Yes	Low	B(NS)	Rigamonti et al. 2005
Homo sapien	72	0.02	10	10	8	32	Yes	No	1	Yes	High	B(Z=2.22,p=.013)	Rigamonti et al. 2005
Homo sapien	73	0.03	10	10	8	32	Yes	No	1	No	High	B(Z=2.22,p=.013)	Rigamonti et al. 2005
Homo sapien	73	0.03	10	10	8	32	Yes	No	1	Yes	Low	B(NS)	Rigamonti et al. 2005
Homo sapien	73	0.03	10	10	8	32	Yes	No	1	Yes	Low	B(NS)	Rigamonti et al. 2005
Macaca nemestrina	70	0.24	7	0	4	4	Yes	No	2	No	Low	NS	Rigamonti et al. 2005
Homo sapien	71	0.01	5	2	4	4	Yes	No	2	No	Low	NS	Rigamonti et al. 2005
Homo sapien	72	0.02	5	5	4	4	Yes	No	2	No	Low	NS	Rigamonti et al. 2005
Homo sapien	73	0.03	5	4	4	4	Yes	No	2	No	Low	NS	Rigamonti et al. 2005
Pan trogoldytes	74	0.28	4	4	3	5	Yes	No	5	Yes	High	KT(p<.03)	Whiten 1998
Varecia variegata	75	0.36	8	8	9.25	223	Yes	No	1	Yes	High	U(Z=-2.31, p=.029)	Stoinski et al. 2011
Cebus apella	79	0.16	11	11	4	4	Yes	No	3	No	High	U(5,6)=3.5, p=.016 U(5,6)=3.3, p=.017	Custance et al. 1999
Cebus apella	79	0.16	11	11	4	4	Yes	No	3	No	Low	U(5,6)=13, p=.355 U(5,6)=12, p=.283	Custance et al. 1999
Homo sapien	81	0.02	8	8	4	4	Yes	No	3	No	High	U(3,4)=0, p=.02	Whiten et al. 1996
Homo sapien	81	0.02	8	8	4	4	Yes	No	3	No	High	U(4,4)=0, p=.01	Whiten et al. 1996
Homo sapien	83	0.03	8	8	4	4	Yes	No	3	No	High	U(4,4)=0, p=.01	Whiten et al. 1996
Homo sapien	83	0.03	8	8	4	4	Yes	No	3	No	High	U(4,4)=0,p=.01	Whiten et al. 1996
Homo sapien	85	0.04	8	8	4	4	Yes	No	3	No	High	U(4,4)=0, p=.01	Whiten et al. 1996
Homo sapien	85	0.04	8	8	4	4	Yes	No	3	No	High	U(4,4)=0, p=.01	Whiten et al. 1996
Pan trogoldytes	87	0.08	7	7	4	4	Yes	No	3	No	High	U(3,4)=0, p=.02	Whiten et al. 1996
Pan trogoldytes	87	0.08	8	8	4	4	Yes	No	3	No	Low	U(4,4)=5,p=.2	Whiten et al. 1996
Saguinus oedipus	89	0.02	13	8	22	19	Yes	No	1	No	Low	FET(p=0.54)	Hulme & Snowdon 2008
Callithrix jacchus	91	0.47	5	5	1	19	Yes	No	1	Yes	Low	U(5,6)=11.5, p>.05	Bugnyar & Huber 1997

Cebus apella	93	0.29	12	12	1	40	Yes	No	1	Yes	Low	U(6,6)=16.5, p=.82	Dindo et al. 2010
Cebus apella	95	0.18	5	4	3	5	Yes	Yes	1	Yes	High	U(3,3)=0, p=.034	Fredman & Whiten 2008
Cebus apella	98	0.16	8	7	3	5	Yes	No	1	Yes	Low	U(4,4)=0, p=.0014 U(4,4)=3, p=.15	Fredman & Whiten 2008
Cebus apella	100	0.20	9	9	1	20	Yes	No	1	Yes	High	U(4,4)=-2.61, p=0.01	Dindo et al. 2008
Pan trogoldytes	104	0.10	10	4	50	50	Yes	Yes	1	No	Low	H(NS)	Nagell et al. 1993
Homo sapien	107	0.02	16	10	10	10	Yes	Yes	1	No	High	F(2,22)=4.48, p<.05	Nagell et al. 1993
Pongo pygmaeus	109	0.27	16	11	50	50	Yes	Yes	1	No	Low	U(NS)	Call & Tomasello 1994
Macaca nemestrina	111	0.21	11	6	4	6	Yes	No	3	No	Low	W(Z=-1.138, p=.128)	Custance et al. (2006)
Homo sapien	114	0.03	22	18	1	2	Yes	Yes	7	No	High	F(2,20)=48.4, p<.001	Whiten et al. 2006
Pongo pygmaeus	116	0.36	9	7	3	5	Yes	No	5	Yes	High	U(4,5)=1, p<.05	Stoinski & Whiten 2003
Pan trogoldytes	118	0.38	12	12	1	10	Yes	No	1	Yes	High	FET (p<0.001)	Horner et al. 2006
Homo sapien	120	0.03	16	16	1	2	Yes	No	1	Yes	High	FET(p<.001)	Horner et al. 2006
Saguinus oedipus	122	0.32	9	6	6	6	Yes	No	1	No	Low	FET(NS)	Dillis et al. 2009
Pan trogoldytes	123	0.38	24	20	1	50	Yes	No	1	Yes	High	U(9,11)=78.5, p=0.02 U(9,11)=27.5, p=0.05	Hopper et al. 2012
Pan trogoldytes	4	0.45	8	3	1	0	NA	No	1	Yes	Control	-	Hopper et al. 2008
Homo sapien	8	0.03	8	6	1	0	NA	No	1	Yes	Control	-	Hopper et al. 2008
Pan trogoldytes	11	0.21	6	5	1	0	NA	No	1	Yes	Control	-	Tennie et al. 2006
Pan trogoldytes	20	0.14	3	2	1	0	NA	No	1	Yes	Control	-	Tennie et al. 2006
Pan paniscus	21	0.17	1	1	1	0	NA	No	1	Yes	Control	-	Tennie et al. 2006
Gorilla g. gorilla	22	0.15	2	2	1	0	NA	No	1	Yes	Control	-	Tennie et al. 2006
Pongo pygmaeus	23	0.14	2	2	1	0	NA	No	1	Yes	Control	-	Tennie et al. 2006
Homo sapien	30	0.01	12	10	1	0	NA	No	1	Yes	Control	-	Tennie et al. 2006
Homo sapien	31	0.01	14	10	1	0	NA	No	1	Yes	Control	-	Tennie et al. 2006
Homo sapien	32	0.02	10	10	1	0	NA	No	1	Yes	Control	-	Tennie et al. 2006
Pan trogoldytes	36	0.19	13	8	1	0	NA	No	1	No	Control	-	Call et al. 2005

Homo sapien	40	0.02	12	6	1	0	NA	No	1	No	Control	-	Call et al. 2005
Callithrix jacchus	43	0.47	4	0	3	0	NA	No	3	Yes	Control	-	Caldwell & Whiten 2004
Callithrix jacchus	46	0.47	25	18	1	0	NA	No	1	Yes	Control	-	Voelkl & Huber 2000
Homo sapien	55	0.03	16	3	1	0	NA	Yes	1	No	Control	-	Hopper et al. 2010
Pan trogoldytes	65	0.44	10	1	3	0	NA	Yes	2	No	Control	-	Price et al. 2009
Pan trogoldytes	66	0.45	10	2	3	0	NA	Yes	2	No	Control	-	Price et al. 2009
Gorilla g. gorilla	69	0.65	2	0	3	0	NA	No	5	Yes	Control	-	Stoinski et al. 2001
Macaca nemestrina	70	0.24	6	0	4	0	NA	No	2	No	Control	-	Rigamonti et al. 2005
Homo sapien	71	0.01	5	0	4	0	NA	No	2	No	Control	-	Rigamonti et al. 2005
Homo sapien	72	0.02	5	4	4	0	NA	No	2	No	Control	-	Rigamonti et al. 2005
Homo sapien	73	0.03	5	4	4	0	NA	No	2	No	Control	-	Rigamonti et al. 2005
Callithrix jacchus	92	0.47	6	2	1	0	NA	No	1	Yes	Control	-	Bugnyar & Huber 1997
Cebus apella	102	0.20	4	3	1	0	NA	No	1	Yes	Control	-	Dindo et al. 2008
Pan trogoldytes	103	0.11	5	1	50	0	NA	Yes	1	No	Control	-	Nagell et al. 1993
Homo sapien	106	0.02	8	2	10	0	NA	Yes	1	No	Control	-	Nagell et al. 1993
Homo sapien	115	0.03	9	0	1	0	NA	Yes	7	No	Control	-	Whiten et al. 2006
Pongo pygmaeus	117	0.29	5	1	3	0	NA	No	5	Yes	Control	-	Stoinski & Whiten 2003
Pan trogoldytes	119	0.47	6	3	1	0	NA	No	1	Yes	Control	-	Horner et al. 2006
Homo sapien	121	0.03	15	8	1	0	NA	No	1	Yes	Control	-	Horner et al. 2006
Pan trogoldytes	124	0.40	12	6	1	0	NA	No	1	Yes	Control	-	Hopper et al. 2012

Key to test statistic abbreviations: B=binomial test; F=anova; FET=Fisher's exact test; H=Kruskal Wallace; KT=Kendall's Tau; NS=non-significant result reported, but details not given; T=t-test; U=Mann-Whitney U test; X^2 = Chi square test.

2.3. Analysis

All analyses were carried out in R version 3.1.3. Analyses were conducted using binomial generalized linear mixed models (GLMMs) in package *lme4* (Bates et al. 2014). The response variable used in analyses was the proportion of a test group that successfully retrieved the reward from the apparatus (i.e., the number of successful versus unsuccessful subjects in a group-level learning assessment), which equated to the measure of *outcome fidelity* when the test group had witnessed a demonstration. Adopting a model with a binomial error structure ensured that each test group was weighted by its relative size (i.e. the number of individual test subjects) during model fitting.

2.3.1. Preliminary analysis

In an initial analysis, a subset of the data, including only those test groups with a matched asocial learning control group, was used to assess the effect of receiving social information on successful reward retrieval. Firstly, I ran the model (*Model 1*) with a binary fixed effect that detailed whether a test group had received a demonstration or was an asocial learning control. This was followed by a second analysis (*Model 2*) where the original fixed effect was further deconstructed to categorise groups that had received a demonstration according to whether demonstrator fidelity was *high* or *low*. In both models, random effects for species and experimental study were included to control for systematic variance occurring between test groups of the same species or from within the same experimental protocol. An additional observation-level random effect was included to counter over-

dispersion in model fit. Main effects were tested using likelihood-ratio tests (LRTs) and posthoc comparisons were performed using the package *multcomp* (Horthorn et al. 2008; 2013). LRTs compare the change in residual deviance (which follows a X^2 distribution) that occurs between nested models following the removal/addition of a variable.

2.3.2. Main analysis

All test-group-level assessments of social learning were included in the main analysis. Asocial learning control groups were excluded. Eight group-level variables and two interaction terms were included in the global model as fixed effects: (1) the inferred demonstrator fidelity (*high* or *low*; binary variable); (2) the measure of task complexity (numerical variable); (3) an interaction term between demonstrator fidelity and task complexity; (4) whether a tool was used (binary variable); (5) an interaction term between tool use and demonstrator fidelity; (6) whether the task was opaque or transparent (binary variable); (7) the total number of demonstrations (numerical variable); (8) the total number of trials (numerical variable); (9) the mean adjusted age of the test group (numerical variable); and, (10) whether the test group received full or reduced social information during the demonstration (binary variable). Species and experiment were included as random effects to control for systematic variation in the data at these levels. An additional random effect for test group was included, to control for autocorrelation in the response resulting from repeated measures on some test groups.

Model selection (main analysis):

Model selection was undertaken only for the fixed effects. Marginality constraints on interaction effects were respected during model selection (i.e., candidate models containing interaction effects always retained the respective lower-order main effects represented in the interaction). Starting with the full global model, backward stepwise elimination of variables was undertaken, using the likelihood-ratio test statistic and AICc as model comparison criteria (variables that generated non-significant test statistics *and* resulted in a reduction in AICc following removal from the model were eliminated). As stepwise model selection has been criticised for not guaranteeing selection of the best combination of predictor variables (Burnham & Anderson, 2002), the *dredge* function in the *MuMIn package* (Barton, 2015) was also used to undertake an exhaustive search of all predictor variable combinations, in order to identify the best possible model using AICc model comparison. In addition, the relative variable importance (computed as the sum of model AICc weights across all models in a set where the variable occurs) was obtained for the top ten models using the function *model.avg*.

2.3.3. Model fit and diagnostics

Models from the main analysis and preliminary analysis were checked for overdispersion and to ensure that the magnitude of the standardized residuals were independent of the fitted values (Bolker et al., 2009). The full set of predictor variables included in the global model were examined with the *vif* function in the *car* package (Fox & Weisberg, 2011) for evidence of multicolinearity, using the convention that a variable with a variance inflation factor (VIF) of >10 is cause for

concern (Quinn & Keough, 2007). All predictor variables included in analyses had a VIF of <3. Model fit was investigated visually via plots of the fitted verses actual values of the dependent variable (i.e., the proportion of successful test-group subjects), and by obtaining the R^2 of the linear relationship between these two values: $R^2_{(FvA)}$.

2.3.4. Representation and robustness of results

Odds ratios were calculated to provide a measure of effect sizes for the fixed effects. For the graphical representation of variable effects, population mean estimates, which are estimates computed at the average values of the other predictor variables in the model, and are therefore adjusted for variation that is attributable to these other predictor variables, were obtained using package *lsmeans* (Lenth, 2015).

All reported models were refitted following the removal of group-level learning assessments that had been flagged during data compilation (N removed learning assessments: demonstration=5; asocial=0). Due to the relatively high representation of human children in the dataset, models were also refitted following the removal of all human test groups (N removed human learning assessments: demonstration= 37; asocial=12). A subset of experiments in the meta-dataset (N= 12) included only subjects that had successfully retrieved the reward in test-group-level assessments of learning fidelity (i.e., test subjects that had interacted with the apparatus but not retrieved the reward were not considered). This risks introducing error if the behaviour of unsuccessful group members would alter the assignment of any such test groups to high- or low-fidelity learning categories. Thus, all reported models were again refitted following the exclusion of all experiments that included

only successful subjects in fidelity assignment (N removed learning assessments: demonstration= 20; asocial=15).

Finally, for a subset of experiments, information was available regarding the success of subjects on their first test trial (T1) (N available T1 learning assessments: demonstration=53; asocial=27). Hence, models were refitted to this subset, with the number of subjects that were successful versus unsuccessful on their first trial as the dependent variable (n.b. in all other analyses, the dependent variable detailed the success/failure of subjects across the entire experimental procedure).

Where robustness checks failed to replicate the results obtained with the full dataset, this is reported.

2.3.5. Phylogenetic signal

Data collected across multiple species might exhibit phylogenetic signal (i.e., the tendency of closely-related species to be more phenotypically similar than species selected at random from a phylogenetic tree: Bloomberg & Garland, 2002, p905). When undertaking analyses using multi-species datasets, phylogenetic statistical methods are preferred when phylogenetic signal is high, to account for the non-independence of data resulting from phylogenetic relatedness (Felsenstein, 1985; Carvalho et al., 2006). However, phylogenetic signal in behavioural data is often low and/or inaccurate; possibly due to measurement error and noise resulting from variation in measurement protocols (Blomberg et al., 2003; Ives et al., 2007). Additionally, previous analytical investigations have indicated that phylogenetic signal cannot be estimated accurately using datasets with <30 species (Freckleton et al., 2002; see also Postma & Charmantier, 2007; de Villemereuil et al., 2013; and

Quinn et al., 2006, for problems relating to small species samples), and that an unequal distribution of data measures across species might be problematic (Struthsaker, 2000). As such, a phylogenetic analysis was not preferred over conventional statistical techniques for the purpose of this study. However, in order to safeguard that a phylogenetic model would not have generated results inconsistent with those resulting from conventional modelling approaches, I also conducted the main analysis within a Bayesian phylogenetic modelling framework, using the package *MCMCglmm* (Hadfield, 2010).

A consensus phylogenetic tree, based on a sample of 10,000 primate phylogenies in proportion to their posterior probabilities, was downloaded from the website 10KTrees (Version 3, Arnold et al. 2010), with branch lengths scaled according to evolutionary time. The global phylogenetic GLMM followed the same specification of fixed and random effects as detailed for the main analysis in section 3.2, but with an additional random effect included to account for phylogenetic signal.

In *MCMCglmm,* priors for the fixed effects, which are normally distributed, are obtained automatically during the model fitting process, while the random effect priors are given an inverse-Wishart distribution (Hadfield, 2010). The model's sensitivity to the priors was investigated using three different prior distribution scenarios for the random effects, which varied the expected (co)variance *V* and the degree of belief parameter *nu* for the inverse-Wishart distribution: V=1 and nu=0.002 (prior 1); V=1 and nu=0.1 (prior 2); V=5 and nu=1 (prior 3). Although the spread of the posterior distribution for the random effects changed across the three different priors, the proportional contribution of each random effect to the overall variance remained relatively constant. The effect sizes of the fixed effects also varied slightly across the three different priors, but not in direction or significance. The prior that

was least informative and represented the lowest degree of belief (i.e., V=1, nu=0.002) was subsequently used during variable selection.

The model was run for 3,000,000 iterations, with a burn-in of 1,000,000 iterations, and thinned every 500 updates, resulting in an effective sample size of 4000 values. In order to check convergence, the model was run with three different starting values that were chosen specifically to be heuristically unfavourable. Autocorrelation between successive samples of the posterior distribution was low for all parameters (<.06) indicating good chain mixing. Chain convergence was checked visually and using Gelman and Rubin's convergence diagnostic (Gelman & Rubin, 1992). Variable selection was undertaken by backward stepwise elimination using the MCMC *p*-values returned by *MCMCglmm*.

Phylogenetic signal was estimated from the phylogenetic variance in the model using the heritability statistic (h^2) (Hadfield 2010). The value of h^2 ranges from 0 to 1 (values close to 0 indicate negligible phylogenetic signal while values close to 1 indicate a strong effect of phylogenetic relatedness), and is conceptually identical to the parameter lambda (λ) in PGLS regression (Freckleton et al. 2002).

2.4. Results

2.4.1. Preliminary analysis: asocial versus social learning

The results of the preliminary analysis are detailed in Table 2.3. As predicted, irrespective of the social learning mechanism used, receiving a demonstration had a highly significant effect on successful learning outcomes relative to asocial learning (see Table 2.3, *Model 1*). Contrasts between *asocial* learning groups and groups

demonstrating *high* or *low* demonstrator fidelity confirmed expectations that, after controlling for differences between species and experimental procedures, *high* relative to *low* demonstrator fidelity afforded the highest levels of outcome fidelity (see *Model 2* & Figure 2.2). In addition, groups exhibiting *low* demonstrator fidelity were still able to outperform groups that were reliant on *asocial* or individual learning processes.

Robustness checks (see above section 2.3.4) revealed no changes in the direction or significance of these results.

Table 2.3. *Results of preliminary analysis: The effect of receiving a demonstration on the proportion of a test group that learned to retrieve the demonstrated reward*

Model parameters	Pairwise comparisons	Estimate (S.E.)	Odds Ratio
Model 1			
Intercept		-0.31(0.36)	
Learning Type[a]		1.54(0.30)***	4.65
$R^2_{(FvA)}$=0.90; *N* groups: Asocial = 31; Demonstration = 49; *N* subjects=833			
Model 2			
Intercept		-0.36(0.34)	
Learning Type[b]	Demo (Low) – Asocial	1.01(0.32)**	2.74
	Demo (High) – Asocial	2.29(0.39)***	9.84
	Demo (High) – Demo (Low)	1.28(0.41)**	3.58
$R^2_{(FvA)}$=0.90; *N* groups: Asocial= 31; Demo(high)= 31; Demo(low)= 18; *N* subjects=833			

Binomial GLMM with random effects for species, experiment and observation-level. Main effects were tested using likelihood-ratio tests. Pairwise comparisons were undertaken using Wald tests with Tukey correction for familywise error.
[a] Dichotomous variable (Asocial learning = 0; demonstration = 1)
[b] Categorical variable (Asocial = asocial learning; Demo (Low) = groups evidence low fidelity to demonstration; Demo (High) = groups evidence high fidelity to demonstration
**p<.01; ***p<.00

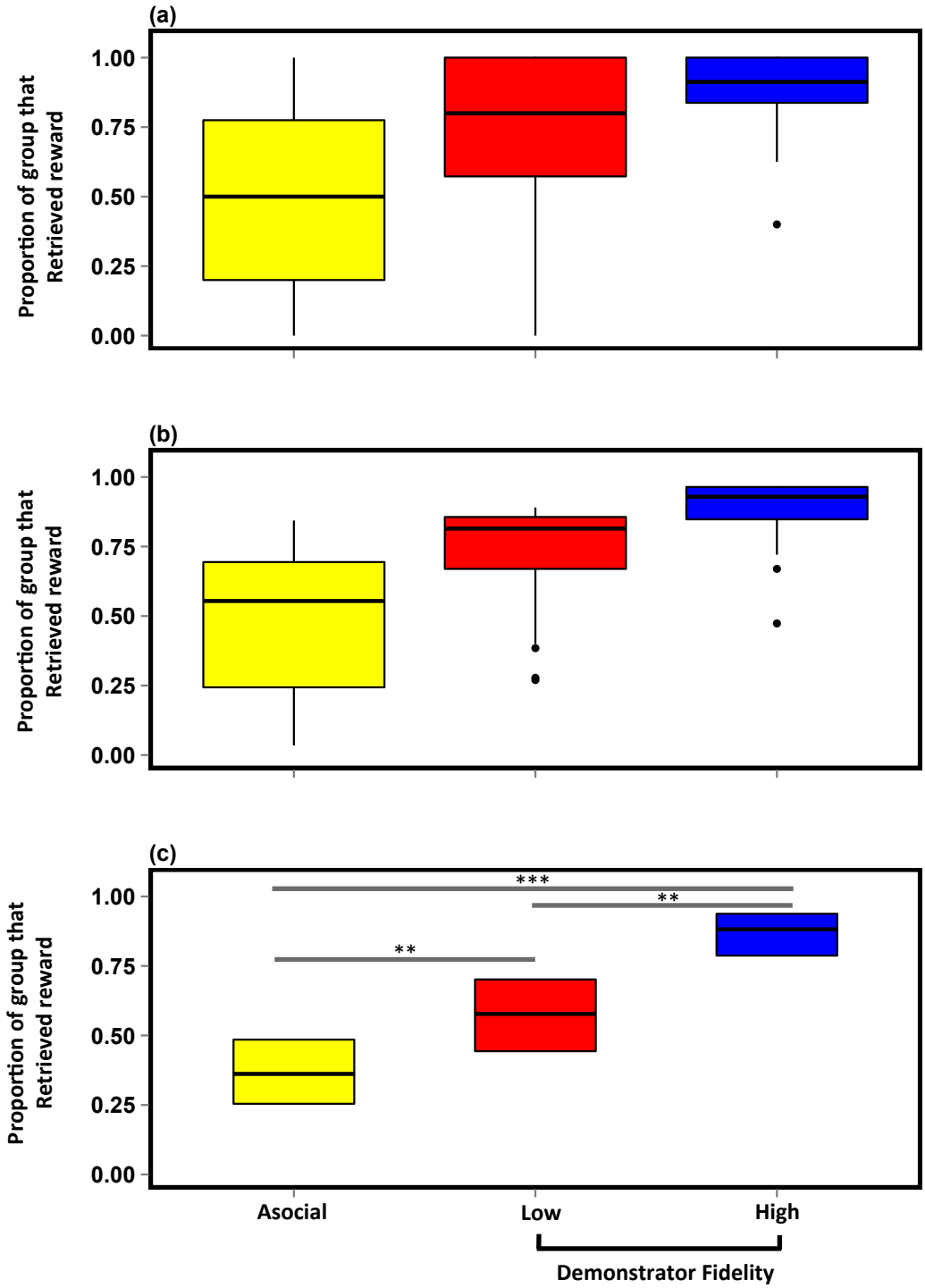


Figure 2.2. *The effect of receiving a demonstration on the proportion of a test group that learned to retrieve the demonstrated reward, at different levels of fidelity (high vs. low) to the demonstrated method (**Model 2, Table 2.3**). Plots represent: (a) boxplot of the raw data values; (b) boxplot of the model-fitted data values; (c) crossbar plot of the model-estimated population means adjusted for sampling variation between learning categories (±95% CI).* $^{**}p<.01$; $^{***}p<.001$

2.4.2. Main analysis: Demonstrator fidelity and task complexity

The main analysis examined the relationship between demonstrator fidelity and outcome fidelity in more detail, including consideration of a number of additional predictor variables on the dependent variable. Model selection undertaken by backward stepwise elimination, using the likelihood-ratio test in combination with AICc, identified the same top model as the exhaustive search of all possible variable combinations based solely on AICc (see Table 2.4 & Table 2.5). The measure of relative variable importance was also concordant with model selection, showing that, with only one exception, all variables retained in the final model were present in all of the top ten AICc ranked models (n.b. lower order terms for the inferred demonstrator fidelity and task complexity were always retained where present in the higher-order interaction term). All predictor variables eliminated during model selection had low relative variable importance (≤0.4). Five predictor variables were included in the final model, the results of which are detailed in Table 2.6.

As predicted, there was a significant interaction between the inferred demonstrator fidelity (*high* or *low*) and task complexity, indicating that *high* demonstrator fidelity (i.e., imitation/emulation) became increasingly important at higher task complexities. A detailed breakdown of the effects of demonstrator fidelity at each level of task complexity is given in Table 2.7 and Figure 2.3. At the lowest level of task complexity – the level adopted in the majority of social learning experiments – there is no significant advantage of learning with *high* demonstrator fidelity on outcome fidelity (i.e., successful learning). However, at all subsequent levels of task difficulty, *high* demonstrator fidelity is associated with significant increases in outcome fidelity, showing increases in the odds ratio at each increment in task complexity between test groups with *high* verses *low* demonstration fidelity.

Tool use, in line with predictions, had a significant and negative effect on learning success; the odds of reward retrieval being approximately ten times lower when using a tool (see Table 2.6 & Figure 2.4: left-panel plots). The expectation that *high* demonstrator fidelity would afford a greater relative advantage in tool-use tasks was not confirmed, with the interaction between demonstrator fidelity and tool use eliminated during model selection, and exhibiting a relative variable importance of 0 (see Table 2.4 & Table 2.5; Figure 2.4: right-panel plots). Figure 2.4(f), which displays the model-estimated effects of tool use at each level of demonstrator fidelity, adjusted for the effects of other predictor variables, suggests a trend in the predicted direction, but also indicates large margins of error due to the relatively small number of tool-use studies (N groups =10) in the sample.

Receiving a complete rather than reduced demonstration of the action- and object-oriented processes required to successfully operate the apparatus appeared to have a marginal positive effect on learning success ($p<.1$; odds ratio=2.07; relative variable importance of 0.61; see Tables 2.4-6). The majority of test groups that received reduced social information during demonstration of the apparatus (e.g., were assigned to 'ghost' conditions), subsequently exhibited evidence of *low* rather than *high* demonstrator fidelity (14/18 test groups = 78%).

Robustness checks using subsets of the data (see above section 2.3.4) revealed no changes in the direction of reported results. Refitting the model with the proportion of individuals who solved on the first trial as the dependent variable, revealed that tool use was no longer a statistically significant predictor of outcome fidelity ($X^2(1)=0.61$, $p=.44$, N groups=53). However, there were only 5 test groups that used a tool in the data subset that detailed success on the first learning trial, substantially reducing the power to detect a significant effect of this variable.

Table 2. 4. *Main analysis: Variable selection undertaken by backward stepwise elimination using the likelihood-ratio test (χ^2) and AICc*

Model parameters	**Variable removed/tested***	**χ^2 (*p*-value)†**	**AICc**
Global model: **DemoFidelity+Complexity+ FullDemo+Tool+TaskOpaque+Age+Ndemos+Ntrials+ DemoFidelity *Complexity + DemoFidelity*Tool**			255.02
DemoFidelity +Complexity+FullDemo+ Tool+Age+Ndemos+Ntrials+ DemoFidelity *Complexity + DemoFidelity*Tool	TaskOpaque	0.07 (*p*=0.79)	252.20
DemoFidelity +Complexity+FullDemo+ Tool+Age+Ndemos+Ntrials+ DemoFidelity *Complexity	DemoFidelity *Tool	0.09 (*p* =0.76)	249.48
DemoFidelity+Complexity+FullDemo+ Tool+Age+Ndemos+ DemoFidelity *Complexity	Ntrials	2.39 (*p* =0.12)	249.15
DemoFidelity +Complexity+FullDemo+ Tool+Age+ DemoFidelity*Complexity	Ndemos	1.68 (*p* =0.19)	248.17
Final model: **DemoFidelity+Complexity+ FullDemo+ Tool+ DemoFidelity *Complexity**	Age	1.40 (*p* =0.24)	246.99
Variables tested after this point were retained in the final model			
DemoFidelity +Complexity+Tool+ DemoFidelity *Complexity	FullDemo	2.94 (*p* <0.09)	247.43
DemoFidelity +Complexity+FullDemo+ DemoFidelity*Complexity	Tool	8.32 (*p* <0.004)	252.81
DemoFidelity +Complexity+Tool+FullDemo	DemoFidelity *Complexity	10.31 (*p* <0.002)	254.79

*Likelihood ratio test *(χ^2)* used to evaluate the effect of removing the specified variable from the most recent model.

† p<.05 indicates that removing the variable significantly reduces the performance of the model.

Table 2.5. *Main analysis: The 10 best supported models and the relative importance of predictor variables included in this candidate set, following an exhaustive search based on AICc*

		AICc top models									
Predictor	Relative variable importance*	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Complexity	1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DemoFidelity	1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Age	0.4			✓	✓			✓		✓	✓
Ndemos	0.34						✓	✓	✓	✓	✓
Ntrials	0.14					✓					✓
TaskOpaque	0.00										
FullDemo	0.61	✓		✓		✓	✓	✓			✓
Tool	1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DemoFidelity *Tool	0.00										
DemoFidelity *Complexity	1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

*Computed as the sum of AICc weights across all the models in the set where the variable occurs.

Table 2.6. *Main analysis: The effects of predictor variables included in the final model on outcome fidelity (proportion of group-level reward retrieval)*

Model parameters	Estimate (S.E.)	Odds ratio
Intercept	3.32(0.80)***	
DemoFidelity[a]	0.76(0.72)	-
Complexity[b]	-0.02(0.24)	-
Tool[c]	-2.42(0.81)**	0.09
FullDemo[d]	0.73(0.42)[†]	2.07
DemoFidelity*Complexity	-1.43(0.44)**	-

$R^2_{(Fitted\ vs.\ Data)}$ =0.88; *N* groups= 84; *N* subjects=926

Binomial GLMM with random effects for species, experiment and test group. All interaction terms were retained with their respective main effects and lower order terms. Variables were tested using likelihood-ratio tests (χ^2) that are detailed in Table 4. (Odds ratio effect sizes are given only for model parameters not included in a second order interaction; interaction effects are detailed separately in **Table 2.7**.)

[a] Dichotomous variable (0 = high, 1 = low)

[b] Numeric variable (number of task manipulations required for reward retrieval)

[c] Dichotomous variable (0 = no tool use required, 1 = tool use required)

[d] Dichotomous variable (0 = received reduced demonstration, 1 = received full demonstration)

[†] $p<.1$; ** $p<.01$; *** $p<.001$

Table 2.7. *Main analysis: The effects of fidelity to the demonstrated method on outcome fidelity (proportion of group-level reward retrieval) at each level of task complexity (obtained from final model estimates detailed in* ***Table 2.6****).*

	Covariate-adjusted contrast	Estimate (S.E.)	Odds ratio
Complexity Level			
1	Demo Fidelity: (High) – (Low)	0.67(0.44)^{NS}	1.96
2	Demo Fidelity: (High) – (Low)	2.11(0.52)***	8.21
3	Demo Fidelity: (High) – (Low)	3.54(0.85)***	34.42
4	Demo Fidelity: (High) – (Low)	4.97(1.26)***	144.18
5	Demo Fidelity: (High) – (Low)	6.40(1.68)***	603.95

^{NS}p=.13; ***p<.001. (These results are presented graphically in **Figure 2.3c**)

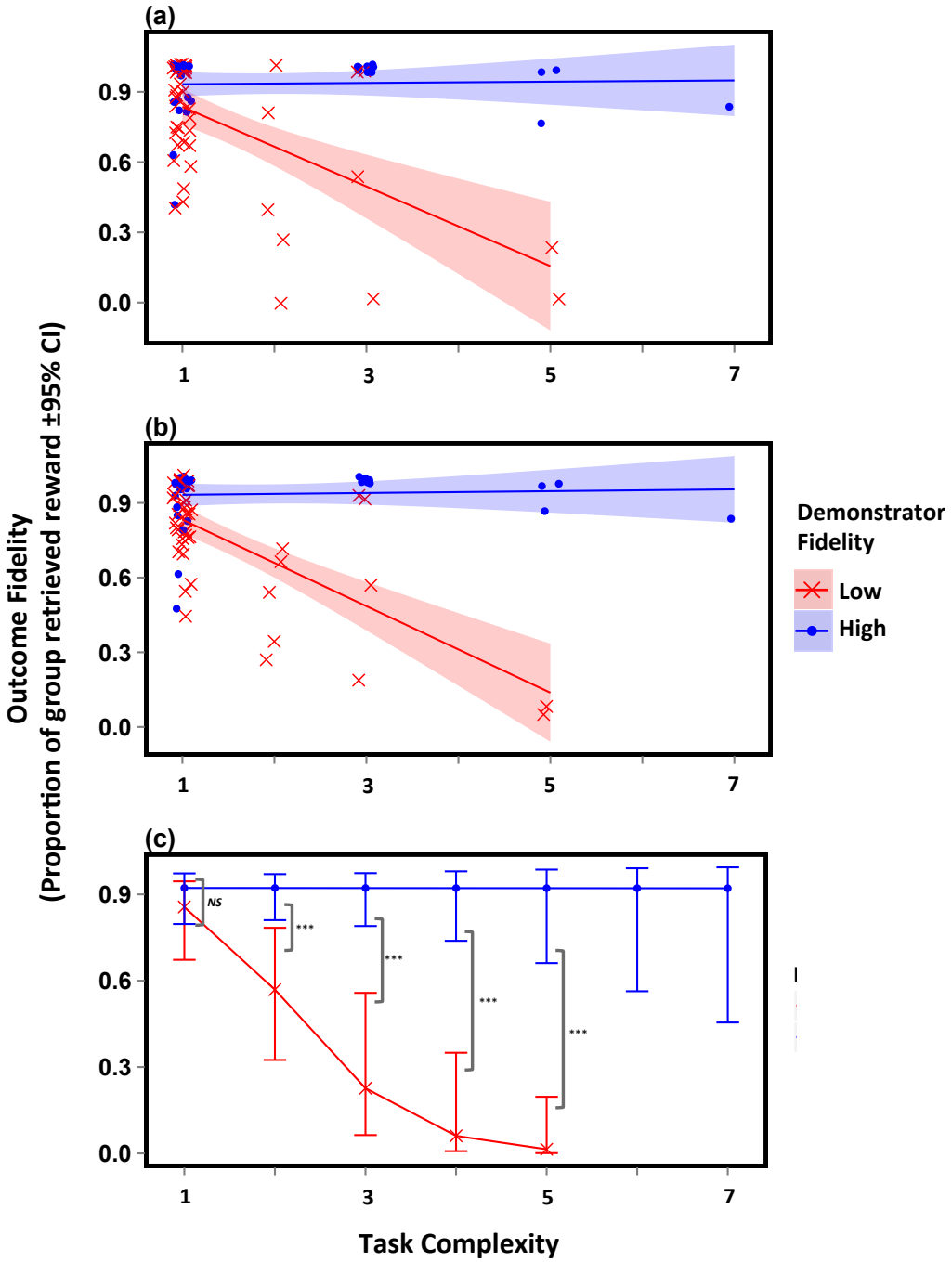


Figure 2.3. *Plots showing the proportion of a test group that learned to retrieve the demonstrated reward (outcome fidelity), at different levels of task complexity and fidelity (high vs. low) to the demonstrated method. Plots represent: (a) raw data values (binomial lines of best fit); (b) model-fitted data values (binomial lines of best fit); (c) model-estimated population mean values (adjusted for the other variables in the final model)*

$^{NS}p>.05$; $^{***}p<.001$

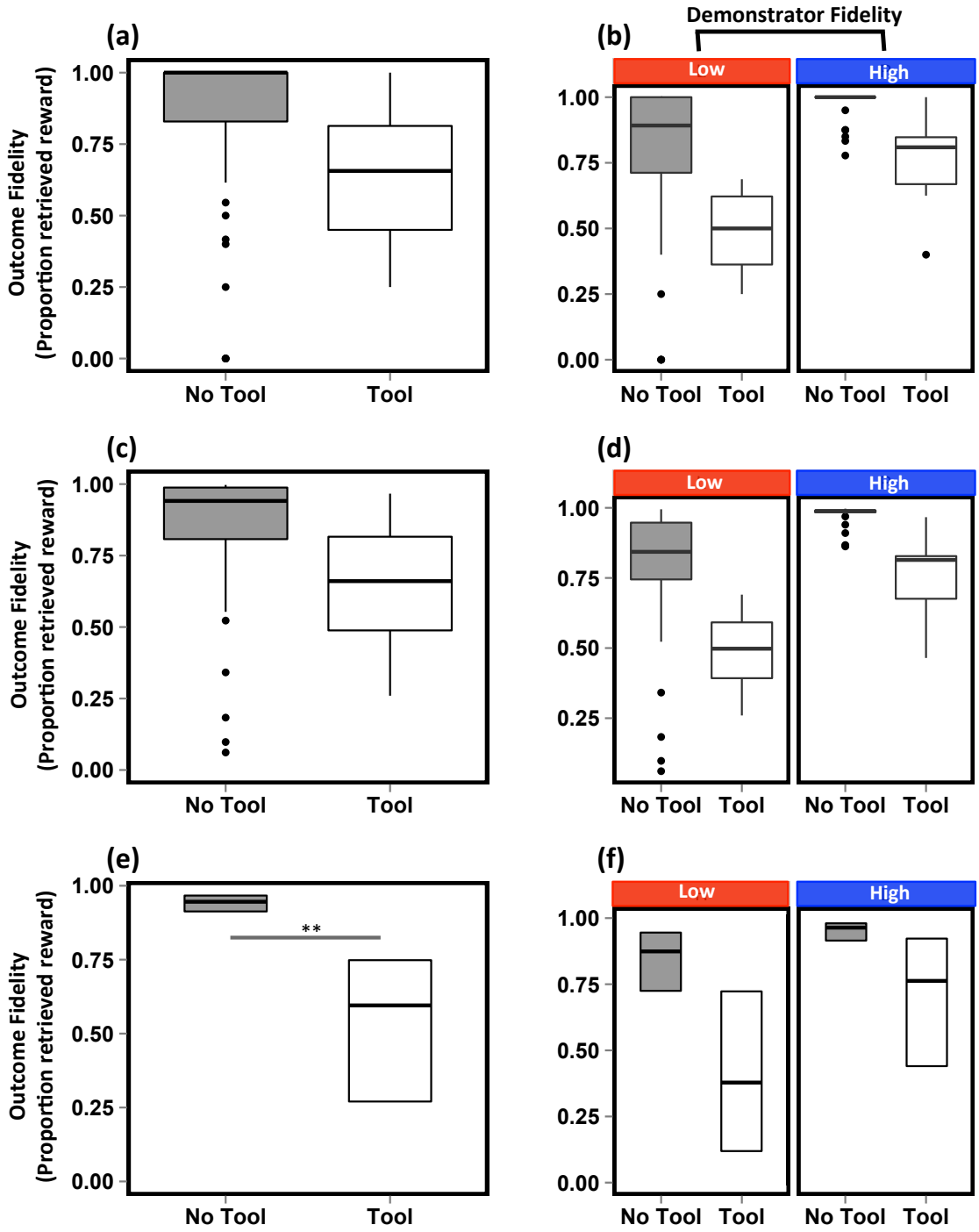


Figure 2.4. *The effect of tool use (left-panel plots) and fidelity to the demonstrated method (right-panel plots) on the proportion of a test group that learned to retrieve the demonstrated reward (outcome fidelity). Plots represent: (a-b) boxplots of the raw data values; (c-d) boxplots of the model-fitted data values; (e-f) crossbar plots of the model-estimated population mean values ±95% CI (adjusted for the other predictor variables in the final model)*

**$p<.01$

2.4.3. Phylogenetic signal

The main analysis was also conducted within a Bayesian phylogenetic framework (see section 2.3.5 for details). A comparison of the final model estimates from the phylogenetic analysis and the conventional non-phylogenetic model detailed in Table 2.6, is summarised in Table 2.8. Both final models contained the same set of predictor variables, and very similar estimates of the magnitude and direction of their effects, with the conventional model providing slightly more conservative estimates. Lambda (λ =0.16, 95% CI=[0.08-0.40]), which was estimated from the total variance in the phylogenetic model, was low, indicating weak phylogenetic signal in these data and the observed concordance between phylogenetic and non-phylogenetic approaches.

Table 2.8. *Comparison of final model estimates generated by a Bayesian phylogenetic GLMM and the conventional non-phylogenetic GLMM in* ***Table 2.6.***

	Phylogenetic model	**Conventional model**
Model parameters	Posterior Estimate [95% CI]	Estimate (S.E.)
Intercept	3.17[0.60 – 5.62]*	3.32(0.80) ***
DemoFidelity[a]	0.77[-0.91 – 2.51]	0.76(0.72)
Complexity[b]	-0.02[-0.52 – 0.48]	-0.02(0.24)
Tool[c]	-2.85[-4.39 – -1.26]***	-2.42(0.81) **
FullDemo[d]	1.02[-0.12 – 2.13] †	0.73(0.42) †
DemoFidelity*Complexity	-1.51[-2.39 – -0.65]***	-1.43(0.44)**

[a] Dichotomous variable (0 = High, 1 = Low)
[b] Numeric variable (number of task manipulations required for reward retrieval)
[c] Dichotomous variable (0 = no tool use required, 1 = tool use required)
[d] Dichotomous variable (0 = received reduced demonstration, 1 = received full demonstration)
† p<.1; *p<.05; **p<.01; ***p<.001 (p-values relate to pMCMC values returned by the package *MCMCglmm,* and likelihood-ratio tests reported in Table 2.4)

2.4. Discussion

The analyses presented in this chapter provide the first extensive review of the relationship between learning mechanisms (demonstrator fidelity) and learning outcomes (outcome fidelity) in primate species. They confirm previous suggestions that learning mechanisms associated with high-fidelity to the demonstrated solution (represented here as a positive result in the two-action task; i.e., imitation/emulation) offer advantages over lower-fidelity mechanisms (e.g., stimulus/local enhancement), in the transmission of instrumental culture, and suggest that the advantages conferred increase incrementally with the complexity of the learning goal. These results support claims that mechanisms affording increased demonstrator fidelity (e.g., imitation/emulation) play a crucial role in the emergence, maintenance and cumulative improvement of culture (e.g., Galef, 1992; Tomasello, 1999; Dean et al., 2012).

Tool use, as expected, increased learning difficulty and exerted a strong negative effect on learning success. The apparent difficulty that was associated with learning to use a tool effectively echoes previous findings (Seed et al., 2009), and might offer some explanation regarding the rarity of tool use in the animal world. Certainly, tool use has been linked with enhanced cognitive abilities (Reader & Laland, 2002; Stout et al., 2008; Reader et al., 2011), and with complex forms of social learning, such as active teaching through language in humans (Stout, 2011; Morgan et al., 2015a). However, the expectation that mechanisms conferring *high* demonstrator fidelity would prove disproportionately advantageous in tool-using tasks was not supported. The relationship between demonstrator fidelity and tool use on learning success, when represented graphically, does suggest a trend in the expected direction. Failure to detect this effect might reflect the relatively low sample of tool-

using studies in the meta-dataset. Alternatively, it is possible that the skills required for effective tool-use in social learning experiments are benefitted by other, more enhanced mechanisms of high-fidelity transmission, such as active instruction and teaching by the demonstrator, or by other skills not connected to social learning, which were not incorporated into these experimental designs.

Test groups that received reduced information about the apparatus manipulations necessary to obtain the goal (e.g., groups assigned to 'ghost' conditions) were less successful than groups that received a full demonstration, but this result was only marginally significant, likely due to the small sample of studies in this category. The majority of test groups receiving reduced information were categorised as exhibiting *low* demonstrator fidelity (or more accurately *low demonstration fidelity* when a demonstrator was not present), suggesting that the negative impact on learning might have resulted from the absence of a social model at the task, which was characteristic of most reduced-information conditions. A study by Hopper et al. (2008), finding that demonstration fidelity and outcomes are enhanced by conspecific presence at the task during 'ghost' demonstrations, supports this suggestion. That most test groups receiving reduced information did not demonstrate *high* demonstration fidelity is also suggestive that information pertaining to both a model's bodily actions, and the associated object movements, is useful in learning instrumental skills, and that a reduction of one or the other leads to a loss of fidelity. Further examination of the relative importance of each type of learning process in instrumental skill learning is an important question for future research (see also Heyes, 2012).

A number of variables relating to variation in inter-experimental designs were eliminated during model selection. It is interesting that variables relating to the

quantity of social information (e.g. number of demonstrations) and the opportunity to be successful (e.g. number of trials) were not influential predictors of learning outcomes. A plausible explanation is that researchers have already calibrated their experimental designs to reflect the optimal local levels of these variables for their test subjects, when the primary goal of their research is to identify the mechanisms that underpin learning outcomes.

The finding that *high* demonstrator fidelity did not afford a significant increase in learning success at the lowest level of task complexity, offers an explanation for previous instances of successful behavioural spread in the absence of mechanisms such as imitation/emulation. It is likely that behaviours such as potato washing in macaques, and milk-bottle-top opening in birds, are simple enough to be transmitted successfully without recourse to copying the detailed actions and object movements. This raises the question of why the majority of social learning experiments have been conducted at the lowest level of task difficulty, when incentives for test subjects to copy using higher-fidelity processes might be reduced. Other authors have also alluded to the possibility that learning processes leading to *high* demonstrator fidelity are superfluous when the task demands are easy (e.g., Byrne & Russon, 1988; Whiten et al. 2009). An interesting avenue of future research will be to investigate whether animals exercise learning mechanisms flexibly in response to the perceived difficulty or uncertainty of success on the task; a heuristic that might complement other proposed learning strategies regarding when and how to use social information (Laland, 2004, Kendal et al., 2005).

However, the possibility also remains that, while imitative or emulative learning did not enhance successful outcomes when tasks were easy, other measures of learning, such as the latency to success, were still favourably impacted. This

suggestion echoes Galef (1992), who argued that imitative learning processes offer a more efficient way of acquiring new behaviours. Yet few studies have provided detailed breakdowns of learning outcomes, which might usefully include measures of latency and the results of subjects' first interaction with the task (though see e.g., Rigamonti et al., 2005 for an exception); possibly reflecting the emphasis on identifying learning mechanisms, while leaving learning outcomes implicitly assumed, which has prevailed in experimental investigations of social learning.

Of course, complex culture also depends on conventional and ritual behaviours, in addition to instrumental skills and the transmission of material culture. A number of authors have suggested that the highest fidelity of human imitation occurs not during the transmission of instrumental culture and artefacts, but rather in learning the ritual and conventional behaviours that signal group identity and affiliation (Over & Carpenter, 2012; Herrmann et al., 2013; Legare & Nielsen, 2015). An expectation that derives from this is that *high* demonstrator fidelity is likely to be advantageous at all levels of ritual and conventional behavioural complexity, where re-enacting the precise actions of other group members is key to preserving group solidarity and homogeneity.

A potential criticism of the current study is the existence of noise in the measure of task complexity. It is possible that some task designs afford manipulations that are easier to execute than others, and are also more amenable to particular species (Caldwell & Whiten, 2002). While this suggestion is plausible, it seems unlikely that adjustments to the task difficulty measure would substantially alter the clear interaction between task difficulty and fidelity found in these data. Another issue that can affect analyses resulting from metadata is the underreporting of null results (known as the 'file drawer problem'), resulting in a biased data sample. However,

given the pattern of reporting demonstrated in the current data, along with the emphasis on identifying learning mechanisms rather than learning outcomes, it seems plausible that unpublished results are likely biased towards studies demonstrating *low* demonstrator fidelity coupled with *low* outcome fidelity, at higher rather than lower levels of task complexity – i.e. data that would enhance the strength of the relationships reported here.

More refined measures than the high- or low-fidelity learning dichotomy offered by the two-action task would also enhance the current findings. Additionally, assessments of demonstrator fidelity and outcome fidelity at the individual level, rather than on the aggregate behaviour of the test group, would enable a more detailed and powerful analysis of the relationships reported here. Systematic, coordinated and controlled comparisons across species, using a range of varying task complexities, and documenting measures of learning fidelity that include and go beyond the typical two-action paradigm (e.g., to include sequential copying (e.g., Whiten, 1998), attribution of intentions (Burkart et al., 2012), latency/efficiency measures and more detailed assessments of action and object copying), would broaden our understanding of high-fidelity social transmission and its evolutionary trajectory across primate species.

The low number of species and over-representation of a few species in these metadata, coupled with inter-study procedural variation, prevented an accurate examination of cross-species variation in learning ability. Previous research by Reader and Laland (2002) and Reader et al. (2011) has indicated that social learning frequencies covary positively with brain size in primates, and has been offered in support of hypotheses claiming that increased reliance on cultural learning has been a key factor driving increases in brain size in this lineage (e.g., Wilson, 1985; Boyd & Richerson, 1985; Henrich & McElreath, 2003; Whiten & van Schaik, 2007).

However, Reader and colleagues' measure of inter-species social learning is comprised of systematic counts of social learning occurrence, as reported in the published primate literature, without recourse to the complexity of social learning presented or the fidelity of its transmission. The results of the current analysis, coupled with observations that simple forms of social learning are common in a variety of small-brained taxa, including invertebrates (Leadbeater & Chittka, 2007), urge caution when inferring relationships between unrefined rates of social learning and brain size.

Moreover, the unequal distribution of cross-species data in the current study implies that published reports of social learning occurrence, both in naturalistic and experimental contexts, are possibly biased towards species *a priori* believed to exhibit enhanced social learning capabilities (i.e. hominids, Cebus and Macaca). Reader and colleagues' attempts to control for research effort have not incorporated this bias, risking circularity in their argument: larger brained species, which are associated with higher levels of intelligent behaviour (including social learning), generate increased reports of social learning because research in social learning is concentrated on these species. A systematic, unbiased and controlled experimental program to assess social learning abilities across a range of primate species, though logistically challenging, would provide invaluable validation to Reader and colleagues' thesis.

In summary, the data and analyses presented here support previous arguments that learning mechanisms such as imitation and emulation play an essential role in the emergence and maintenance of complex cultural behaviour. Given that the extraordinary cultural achievements of human beings have driven our species' global dominance and success, it is highly likely that mechanisms affording high

demonstrator fidelity have undergone increased selection and refinement during the evolution of our lineage. Future studies that generate more refined measures of demonstrator fidelity and outcome fidelity, across different learning complexities and contexts, are required for a more comprehensive understanding of how learning mechanisms influence learning outcomes, and the selective processes that may have favoured their evolution. An exciting and plausible possibility, given the enormous cultural success of the human lineage, is that social learning, and the behavioural advantages that its enhanced forms can attain, has played an important role in driving increases in brain size and intelligence across the primate lineage (Wilson, 1985; Whiten & van Schaik, 2007; Reader et al., 2011). While causality in brain-behaviour correlations is always very difficult to infer (Healy & Rowe, 2007), more refined and systematic measures of social learning abilities across species would clarify the position of previous arguments.

Chapter 3

Selective Copying and 'Over-Imitation': An Experimental Investigation in Young Children

Abstract

The previous chapter examined the social learning mechanisms thought to be involved in the successful high-fidelity transmission of instrumental cultural goals (i.e., imitation/emulation). The current chapter examines children's tendency, during experimental tests of social learning, to copy even those actions that are manifestly causally irrelevant to achieving an instrumental goal – a puzzling phenomenon described as 'over-imitation'. Taking a cultural evolutionary approach, I investigated whether majority-biased copying in children extends to majorities that perform irrelevant actions. I found that children always displayed majority-biased copying when the majority performed an efficient task solution, but that majority-biased copying did not extend to majorities who performed an irrelevant action. Additionally, results indicated that children's copying of the irrelevant action was influenced not by causal inferences but rather by demonstrator behaviour (i.e., socially driven motivations). I discuss whether the term 'over-imitation' is appropriate in instances where causally irrelevant actions encompass socially functional properties. I further argue that, in more real-world contexts, where actions are not always demonstrated unanimously, causally irrelevant actions might substantially alter the operation of adaptive learning biases, and are likely to be quickly eliminated in purely causal or instrumental learning contexts.

3.1. Introduction

Humans show an exceptional ability to learn through copying others. The tendency of young children to engage in high-fidelity copying of others' actions appears to set them apart from other animals (e.g., Dean et al., 2012). This propensity to engage in faithful copying is thought to play a crucial role in facilitating cumulative cultural improvement: a hallmark of human culture (Dean et al., 2014; Tomasello, 1999).

However, children's high-fidelity copying in laboratory studies has also been described as 'surprisingly unselective' or 'mindless' (Whiten et al., 2009), and susceptible to behavioural 'inefficiency' or 'cost' (Lyons et al., 2007; McGuigan et al., 2007), following numerous reports that children blanket copy even those parts of an action sequence that are manifestly causally irrelevant to obtaining the instrumental goal (e.g., Horner & Whiten, 2005; Kenward et al., 2011; Lyons et al., 2007; McGuigan et al., 2007, 2011; Nielsen & Tomaselli, 2010). This phenomenon, which has been dubbed 'over-imitation' (Lyons et al., 2007) and which has received much attention in recent years, has been replicated in several cultures (Nielsen & Tomaselli, 2010; Nielsen et al., 2015), and, surprisingly, is reported to increase with age into adulthood (McGuigan et al., 2011, 2007; Nielsen & Tomaselli, 2010; Whiten et al., 2009).

The seemingly counterintuitive nature of 'over-imitation' has led some authors to propose explanations that are grounded in causal cognition, suggesting that the demonstration leads children to imitate the action automatically (Lyons et al., 2007; 2011; Horner & Whiten, 2005; Whiten et al., 2009), despite being able to understand the causal mechanisms necessary to operate the apparatus. Such blanket copying, it is argued, might serve to promote facets of cultural learning that are causally opaque (Lyons et al., 2007; 2011), and should lead mostly to gains in cultural

knowledge, assuming that the demonstrator is a more experienced adult (Whiten, Horner, & Marshall-Pescini, 2005; Whiten et al., 2009).

Critics of causal explanations have argued that the phenomenon results from more purely social processes, rather than mechanisms grounded in causal understanding (Legare et al., 2015; Nielsen & Blank, 2011; Over & Carpenter, 2012). On this view, the term 'over-imitation' is misleading when applied to contexts in which the causally irrelevant actions encompass socially relevant pressures and benefits: rather, 'optimal-imitation' might be a better term there. For example, children might be motivated to copy the causally irrelevant actions in order to be like, and share experiences with, the demonstrator, or to affiliate with and encourage the demonstrator to like them (Meltzoff, 2007; Nielsen, 2008; Nielsen & Blank, 2011; Over & Carpenter, 2012). Likewise, unanimity in the demonstration of irrelevant actions, which are often provided in a pedagogical context, might additionally lead children to believe that they are expected by the experimenter to perform the irrelevant action (Lyons et al., 2011), or that the demonstration is normative, and that they ought to conform to its performance, despite its function being unclear (Haun & Tomasello, 2011; Kenward et al., 2011; Kenward, 2012; Keupp et al., 2013; Keupp et al., 2015).

The critiques levelled at causal hypotheses resonate with other findings that children's imitation can be selective and strategic. There is compelling evidence that children are able to imitate rationally, adjusting their imitative fidelity flexibly in response to a number of contextual cues, including the demonstrator's competency (e.g., Birch et al., 2008; Williamson et al., 2008) and intentionality (Carpenter et al., 1998), whether the demonstrator was constrained (Gergely et al., 2002), signs of

pedagogical engagement (Csibra & Gergely, 2006), and the perceived goal of the task (Carpenter et al., 2005; Williamson et al., 2008).

Here I take a cultural evolutionary approach to investigating the question of whether children are better characterized as 'over-' or 'optimal-' imitators. Cultural evolutionary theory predicts that social learning decisions should be strategic and selective regarding *whom* and *when* we copy (Boyd & Richerson, 1985), and guided by a set of adaptive learning heuristics that influence the emergence of socially transmitted behaviours within cultural groups (Boyd & Richerson, 1985; Laland, 2004; Rendell et al., 2011b). Evidence that learning biases are involved in guiding our use of social information has been provided using both theoretical (e.g., Boyd & Richerson, 1985; Cavalli-Sforza & Feldman, 1981; Enquist & Ghirlanda, 2007; Kandler & Laland, 2013) and empirical approaches (e.g., Wood et al., 2012; Morgan et al., 2011; Toelch et al., 2014; Mesoudi, 2008), yet have been little considered in investigations of 'over-imitation'. Initial evidence that learning biases are operational in the adoption of causally irrelevant actions suggests that these actions are more likely to be copied when there are adult rather than child models (Flynn, 2008; Wood et al., 2012), or high rather than low status individuals (McGuigan, 2013).

Decisions regarding the adoption of causally irrelevant actions should be especially tuned to adaptive decision-making processes. Most previous 'over-imitation' research has been conducted using one-to-one or otherwise unanimous interactions between demonstrator and observer (although see McGuigan & Robertson, 2015; Nielsen & Blank, 2011 for exceptions), yet 'real world' learning might involve observations of multiple individuals who behave in different ways. Comparing the operation of learning biases across situations that include, exclude, or otherwise vary the degree of irrelevant action performance in the demonstrations will

be particularly informative regarding the robustness of children's propensity to 'over-imitate' outside of these standard unanimous conditions. This chapter addresses this, considering one type of learning bias that has fascinated cultural evolutionists and psychologists alike: majority-biased copying.

There is empirical evidence that majority or consensus behaviour informs copying in both children (Corriveau et al., 2009; Corriveau & Harris, 2010; Fusaro & Harris, 2008; Haun & Tomasello, 2011; Haun et al., 2012; Herrmann et al., 2013; Morgan et al., 2015a) and adults (Coultas, 2004; Morgan et al., 2011). For example, Haun et al. (2012) found that 2-year-old children were more likely to drop balls into a reward-releasing container that was previously chosen by three different individuals once each (majority group) than an alternative container that was chosen by one individual three times (minority person). Because the majority behaviour represents the behaviour that the greatest proportion of group members have converged upon, it is expected to signal a relatively safe, reliable, and adaptive behavioural response (Boyd & Richerson, 1985; Henrich & Boyd, 1998; Hastie & Kameda, 2005; Wolf et al., 2013). It therefore presents a particularly suitable learning bias for testing hypotheses about the adoption of causally irrelevant information.

In the current study, I showed 4- to 6-year-old children a video demonstration in which the number of demonstrators who performed a causally irrelevant action (along with a causally relevant one) while getting a reward out of a puzzle box was systematically varied. That is, either all four of the demonstrators, or the majority (3 out of 4), or the minority (1 out of 4), or none of the demonstrators, performed the causally irrelevant action. Outcome-oriented language emphasizing the instrumental end-goal of the task was used, as previous research indicates that language cues can

influence children's perceptions of the task goal (Herrmann et al., 2013; Legare et al., 2015; Clegg & Legare, 2015).

The first experimental condition examined whether children were more likely to adopt the majority over the minority solution when faced with alternative, but equivalent, *causally relevant* task solutions. In line with previous findings (Haun et al., 2012), I anticipated that children would demonstrate a bias towards copying the majority's solution. Importantly, I then asked whether majority-biased copying in children extends to majorities who perform a causally irrelevant action.

If children copy actions blindly, without regard to their causal efficacy – if they [truly] 'over-imitate' – they should copy the majority regardless of whether the majority solution omits or includes causally irrelevant actions. However, I predicted that they would not do this, and instead that when presented with a majority performing the irrelevant action and a minority omitting it, the instrumental framing of our task, coupled with children's tendency towards rational and selective imitation (e.g., Gergely et al., 2002; Want & Harris, 2001), would counter their tendency to copy the majority, and majority-biased copying would not be detected. In contrast, in a condition in which the majority *omitted* the irrelevant action and the minority performed it, majority-biased copying was expected.

I compared these results to those from a condition representing the paradigm typically used in 'over-imitation' research: unanimous demonstrations of the irrelevant action. Here I predicted that the unanimity of the demonstration would result in irrelevant action copying at similarly high levels as previously reported (e.g., Horner & Whiten, 2005; Lyons et al., 2007). A final condition, with no demonstration, provided the baseline level of irrelevant action production. Participants were provided with multiple (three) attempts at solving the puzzle box, as this

permitted an evaluation both of children's initial tendency to copy as well as their tendency to 'stick with' performing the demonstrated actions after their own initial experience with the task.

I tested 4- to 6-year-olds, as children within this age range have developed sensitivity to demonstrator frequency in copying decisions (Haun & Tomasello, 2011; Haun et al., 2012; Morgan et al., 2015a; Wilks et al., 2015), as well as an ability to engage in rational and selective imitation (Gergely et al., 2002; Want & Harris, 2001). Children in this age range are also considered prolific 'over-imitators' (Kenward et al., 2012; Lyons et al., 2007; Nielsen & Tomaselli, 2010).

3.2. Methods

3.2.1. Participants

Two hundred and fifty-two 4- to 6-year-old children (128 males; 4-year-olds: M = 4;5, range = 4;0 - 4;11; 5-year-olds: M = 5;6, range = 5;0 - 5;11; 6-year-olds: M = 6;5, range = 6;0 - 6;11) were included in the final sample. Eight additional children were tested but were excluded from analyses due to experimenter error (2), apparatus failure (3), parental interference during testing (2), and refusal to interact with the apparatus (1). Participants were recruited at science centres in central Scotland and NE England. Testing took place on non-school days when children were visiting the science centre accompanied by a parent or guardian. Recruitment was via advertisements and visitor information displayed within each science centre, and participation was voluntary. Ethical approval for this study was given by UTREC of the University of St Andrews.

3.2.2. Apparatus

The 'Sweep-Drawer Box' (Wood et al., 2013, see Figure 3.1a-c), a two-action transparent apparatus, was used with minor modifications. A capsule containing a sticker reward can be inserted into the apparatus through a cylindrical inlet located on the top of the box. Once inserted, the reward capsule rests on an opaque mid-level platform, and retrieval is dependent upon the capsule being first moved to the lower level by one of two spatially separated and functionally independent manipulandi: i) a silver sweeper with blue handle (Figure 3.1a), or ii) a blue drawer with red handle (Figure 3.1b). Capsule retrieval using the 'sweep' method involves moving the sweeper towards the front of the box, thereby pushing the capsule through a hole leading to the lower level. Capsule retrieval using the 'drawer' method involves pulling the drawer outwards, which creates a hole in the mid-level through which the capsule falls to the level below. Once at the lower level, the capsule can be retrieved by sliding open the black opaque door (unlike Wood et al., 2013, where the door could slide or lift open).

In some demonstrations, capsule retrieval via the sweep or drawer manipulandum was preceded by an irrelevant (i.e., causally unnecessary) action (see Figure 3.1c). The irrelevant action involved the demonstrator twice sliding the black door open and closed, before the sweep/drawer manipulandum was used.

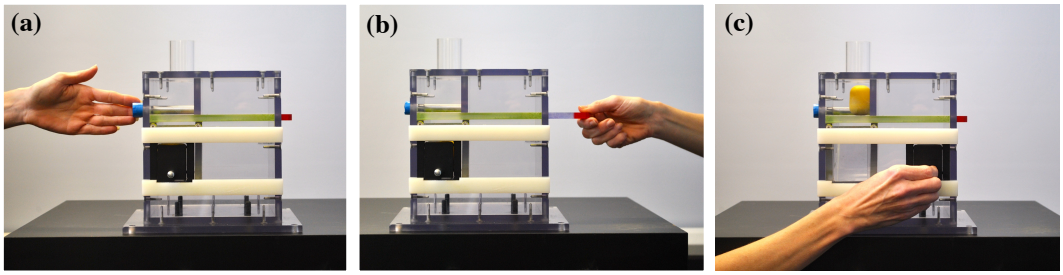


Figure 3.1a – c. *The Sweep-Drawer Box. Demonstrator releasing the capsule using the sweep manipulandum (a), and the drawer manipulandum (b). Demonstrator performing the irrelevant action on the door prior to capsule release (c).*

3.2.3. Design

In a between-groups design, participants were randomly allocated to one of five conditions (C1-C5). There were no significant differences in the distribution of age ($F(4, 246)=0.26$, $p=.91$) and approximately equal numbers of boys and girls in each condition. In four experimental conditions (C1-C4, N=201), children first watched a video showing four different demonstrators retrieving the sticker capsule from the apparatus in turn, before being given their own attempt at capsule retrieval. The fifth condition (C5, N=51) served as a non-social, baseline control in which participants received no video demonstration. The four demonstrators in the video were female, of similar age, and unknown to participants. Each wore a different-coloured long-sleeved t-shirt (red, blue, green, and purple) to help highlight that they were different individuals.

The first experimental condition (*relevant actions only*: C1), in which no irrelevant actions were demonstrated, investigated whether children displayed majority-biased copying when choosing between two causally relevant actions: sweep versus drawer retrieval. Children in this condition saw the majority (three

demonstrators) perform the alternate relevant action to the minority person. In the remaining three experimental conditions, each child saw all four demonstrators perform the same causally relevant action (i.e., sweep *or* drawer), but the number of demonstrators who additionally performed the irrelevant action varied between one (i.e., *minority irrelevant*: C2), three (i.e., *majority irrelevant:* C3), and four (i.e., *all irrelevant*: C4) across conditions. The identity of the minority demonstrator, the order in which the minority and majority performed, and the use of the sweep and drawer methods of retrieval were all counterbalanced within and between experimental conditions (please see Table 3.1 for a detailed outline of the experimental conditions).

The majority demonstrators always appeared consecutively, as a block, with the minority individual demonstrating her method immediately before or immediately after the majority. To control for effects of demonstration frequency, the three majority demonstrators retrieved the capsule once each, while the minority individual demonstrated her method three times. In the *baseline* condition (C5), participants were presented with the apparatus without a prior video demonstration. Each participant received three response trials.

Table 3.1. *Overview of the demonstration and baseline conditions*

Condition	Majority solution (3 demonstrators)	Minority solution (1 demonstrator)	N
(C1) Causal actions only	All retrieve using the same relevant action (sweep OR drawer)	Retrieves using the alternate relevant action	51
(C2) Minority irrelevant	All retrieve using the same relevant action (sweep OR drawer) only	Performs irrelevant action then retrieves using the same relevant action as the majority	51
(C3) Majority irrelevant	All perform the irrelevant action before retrieval. All use the same relevant action (sweep OR drawer)	Retrieves using the same relevant action as the majority, without performing the irrelevant action	49
(C4) All irrelevant	All demonstrators perform the irrelevant action before retrieval. All use the same relevant action (sweep OR drawer)		50
(C5) Baseline	No demonstration		51

3.2.4. Procedure

Children were tested individually in a screened–off area or in a learning classroom at the science centre. Parents were permitted to accompany children during testing, but were seated away from the main test area and asked to refrain from interaction with children until the test session ended. Each child was asked to choose a sticker from a selection, which the experimenter then placed inside the reward capsule. The child was told, "I am going to put this [the capsule] inside a box. All you have to do is try to get it back out of the box. Once you get it out, you can keep the sticker." For the

experimental conditions (C1-4), the child was then shown a picture of the four demonstrators and asked, "Can you see these four people who are all wearing different coloured t-shirts? Earlier I asked each of these four people to take a turn at getting the sticker out, and I made a video of what they did. This is the video of them taking a turn." The child then sat at a table in front of a laptop display screen, with the apparatus visible on a separate table perpendicular to the child. The experimenter dropped the capsule into the top of the box and then started the video.

All video demonstrations began with a still image of the four demonstrators together, accompanied by the video narration: "Earlier I showed the box to four people and asked them to take a turn at getting the sticker out. Watch closely to see what they did." Each demonstrator was then introduced consecutively, and she performed her method of capsule retrieval before the next demonstrator was introduced. The introductions were identical across experimental conditions and consisted of footage of the demonstrator waving and smiling, accompanied by the narration: "Here's the person wearing the (e.g., blue) top. This is her taking a turn." The demonstrator then appeared standing next to the apparatus, ready to take her turn. Video clips of the actual capsule retrieval showed only the demonstrator's arms and hands (Wood et al., 2012), eliminating possible differences in ostensive cues between demonstrators, while still displaying the demonstrators' different coloured shirts. Each retrieval clip ended when the demonstrator had retrieved the capsule and held it in her hand. The footage of the minority demonstrator retrieving the capsule was repeated three times, but she was only introduced once.

After all four demonstrators had retrieved the capsule, the video footage returned to the initial still image of all four demonstrators. The experimenter then verified that the child had watched the video. Pointing at the screen the experimenter

asked, “Did they all get the sticker out of the box?” Only two children answered incorrectly (i.e., “no” instead of “yes”), but subsequently answered correctly after watching the video a second time. The child was next asked to stand in front of the apparatus and told, “Now it’s your turn to try to get the sticker out.” Children were then free to approach the apparatus and interact with it until either (i) the capsule had been retrieved, or (ii) 2 minutes had elapsed, or (iii) the child had refused to continue. When necessary, children were further encouraged, “You can try to get the sticker out if you like, you won’t break the box.” On successful retrieval of the capsule, the experimenter removed the sticker and said, “That’s one sticker for you,” before placing the sticker aside on a nearby table until the test session ended. Participants who successfully retrieved the sticker at T1 were offered two further attempts at sticker retrieval (T2 and T3); the experimenter reset the apparatus out of sight before each subsequent trial while the child was choosing a new sticker.

Children assigned to the *baseline* condition (C5) received the same initial instructions and prompts as children in the experimental groups but watched no video. Before their attempts the experimenter inserted the capsule into the box and reminded the child, “All you have to do is try to get the sticker back out.” All children who participated in the study received a sticker reward.

3.2.5. Coding and inter-observer reliability

Each participant was scored for three separate measures on each response trial: (i) successful removal of the capsule containing the sticker, (ii) the number of times they performed the irrelevant action (sliding the door open and closed prior to operating the manipulandi), and (iii) the manipulandum used during retrieval (sweep or drawer). The experimenter coded 100% of the sample from video records. An independent

observer, who was blind to experimental condition and hypotheses, coded a random sample of 25%. Inter-observer reliability was excellent: Chronbach's alpha = 0.99 for the number of irrelevant actions performed, and Cohen's kappa = 1.00 for the two other measures.

3.3. Analysis

All analyses were carried out in R version 3.1.3. GLMMs and post-hoc comparisons were performed using the packages *lme4* (Bates et al., 2014) and *multcomp* (Hothorn et al., 2008), respectively. Significance testing on regression variables was undertaken by likelihood-ratio tests (LRTs) using the *anova* function. LRTs compare the change in residual deviance (which follows a X^2 distribution) that occurs between nested models following the removal/addition of a variable. All models were checked for evidence of overdispersion, and to ensure that the magnitude of the standardized residuals were independent of the fitted values (Bolker et al., 2009). Two-tailed *p* values are reported throughout.

To assess whether children demonstrated majority-biased copying across all response trials combined (T1-T3), I used the option-bias method (Kendal et al., 2009b), to account for within-individual correlations in responses. This comprised performing a one-sample *t*-test on the difference in the number of times the majority versus the minority was copied within individuals, and then comparing the resulting test statistic (*t*) to the expected null distribution when no preference for the majority or minority solution was observed. The null distribution was computed by randomizing the observed data, recalculating *t*, then repeating this 10,000 times. The null hypothesis could be rejected when the probability (*p*) that the randomization

procedure generated a *t* value at least as big as that calculated for the original data was <0.05.

3.4. Results

Children who received a social demonstration (C1-4) were significantly more successful at retrieving the reward at T1 (success rate = 100%) than those who did not (six participants failed to retrieve the reward in the *baseline* condition: success rate = 88.2%; Fisher's Exact Test, $p<0.001$). All but three participants who retrieved the reward at T1 also retrieved the reward in T2 and T3.

The remaining results are presented in three sections. First, I examine children's copying of unanimous demonstrators. Next the influence of the majority on children's tendency to copy is investigated. Finally, I examine the effect of demonstrator unanimity on children's initial decisions to copy, and children's persistence across trials with the method they first copied.

3.4.1. Copying when the demonstrators were unanimous

Preliminary analyses revealed no effects of children's age or sex on whether children copied actions that were unanimously demonstrated.

Unanimous demonstrations of causally relevant actions

In order to examine the effect of unanimous demonstrators on copying causally relevant information, data was pooled across the three conditions in which children saw all four demonstrators performing the same causally relevant action (i.e., sweep

or drawer retrieval, C2-C4 combined, N=150). Despite successful children in the *baseline* condition showing a bias towards retrieval using the sweep manipulandum (78% of all retrievals used sweep: Fisher's Exact Test, $p<.001$), children who saw a unanimous demonstration showed a strong tendency to copy the relevant action they had witnessed (92% copying across all trials combined [91% sweep, 93% drawer]: Fisher's Exact Test, $p<.001$).

A logistic generalized linear mixed model (GLMM), with standard errors adjusted for repeated measures on participants across trials (see Table 3.2), revealed no significant effect of the method demonstrated (sweep vs. drawer), experimental condition (C2-C4), trial number, or whether the child copied the irrelevant action, on whether the relevant action was copied.

Table 3.2. *The effects of experimental condition, the relevant action witnessed (sweep or drawer), whether the child copied the irrelevant action, and the trial number, on whether the relevant action was copied (C2 - C4)*

Model parameters	Pairwise comparisons	Estimate (S.E.)	Odds ratio
Intercept		10.34(2.26)***	
	Minority (C2) – Majority (C3)	-0.20(2.23)^{NS}	0.82
Condition[a]	*All (C4) – Minority (C2)*	-0.06(2.39)^{NS}	0.95
	All (C4) – Majority (C3)	-0.26(2.52)^{NS}	0.77
Relevant action[b]		0.51(1.90)^{NS}	1.67
Copied irrelevant[c]		0.98(1.34)^{NS}	2.67
Trial[d]		-0.26(0.52)^{NS}	0.77

Logistic generalized linear mixed model (GLMM) with standard errors adjusted for repeated measures on participants. Pairwise comparisons were undertaken with Tukey correction for familywise error. $R^2_{GLMM} = 0.83$ (Nakagawa & Schielzeth, 2013).
[a] Categorical variable: (*Minority irrelevant* (C2) = only one of four demonstrators performs irrelevant action; *Majority irrelevant* (C3)= three of four demonstrators perform irrelevant action; *All irrelevant* (C4) = all four demonstrators perform irrelevant action)
[b] Dichotomous variable (0 = drawer, 1 = sweep)
[c] Dichotomous variable (0 = did not copy irrelevant action, 1=copied irrelevant action)
[d] Numeric variable (trial number)
^{NS} $p>.05$; $^{***}p<.001$

Unanimous demonstration of irrelevant action

Only 16% of children in the *baseline* condition performed the irrelevant action (sliding the door before using the sweep/drawer to move the capsule to the door-level) on their first retrieval attempt (T1). By contrast, a significantly larger percentage of children in the *all irrelevant* condition (C4) copied the irrelevant action at T1 (86%; $\chi^2(1) = 51.60$, $p<.001$), consistent with the high levels of irrelevant action copying reported in previous studies (e.g., Lyons et al., 2007; Horner & Whiten, 2005). Across all trials combined (T1-T3), the percentage of children's responses in the *all irrelevant* condition that included production of the irrelevant action (81%) was significantly greater than in the baseline (9%; $\chi^2(1) = 167.83$, $p<.001$).

3.4.2. Majority-biased copying

Preliminary analyses revealed no effects of children's age, sex, or the method demonstrated (sweep/drawer), on whether they copied the majority behaviour. The order in which the minority and majority performed did influence children's copying, but only in the *majority irrelevant* condition (C3; see below).

Majority-biased copying of causally relevant actions

Figure 3.2 demonstrates that on their first trial (T1) children in the *relevant actions only* condition (C1) copied the majority significantly above chance when faced with demonstrations of two different, yet causally equivalent, relevant actions (binomial test: 76% copied majority, ±95% CI [62% – 87%], $p<.001$). A randomization test (Kendal et al., 2009) revealed that children in this condition continued to demonstrate majority-biased copying when responses across all trials (T1-T3) were considered overall (test statistic = 4.39; $p<.001$; majority: 73%, minority: 27%; see Figure 3.2).

Majority-biased copying of the causally irrelevant action

Participants were scored as demonstrating a majority bias if they copied the majority's behaviour with regard to omitting (*minority irrelevant*; C2) or performing (*majority irrelevant*; C3) the irrelevant action. Figure 2 demonstrates that a strong preference for the efficient majority solution was observed in the *minority irrelevant* condition at T1 (binomial test: 84% copied the majority, 95% CI [71%, 93%], $p<.001$), and remained when responses across all trials (T1-T3) were considered (randomization test statistic = 7.70; $p<.001$; majority: 85%, minority: 15%; see Figure 2).

In contrast, majority-biased copying was not observed in the *majority irrelevant* condition at T1, where the majority of children copied the more efficient minority (binomial test: 41% copied the majority, 95% CI [27% – 56%], $p=.25$). Majority-biased copying was also not observed when responses across all trials (T1-T3) were considered; rather children in this condition showed a bias towards the minority person's more efficient solution that approached significance (randomization test statistic = 1.82; $p<.08$; majority: 39.5%, minority: 60.5%). Children in the *majority irrelevant* condition were influenced by the order in which the majority and minority performed: They more often copied the demonstration they had witnessed first (64% of all responses matched the solution demonstrated first: Fisher's Exact Test, $p<0.001$). Thus, majority-biased copying was always detected when the majority performed an efficient solution, but did not apply to majorities who performed irrelevant actions. Interestingly, 12 children (8%) who had witnessed at least one demonstrator performing the irrelevant action subsequently performed the irrelevant action in a later trial (T2 or T3) after having omitted it at T1 (C2-C4, N=150).

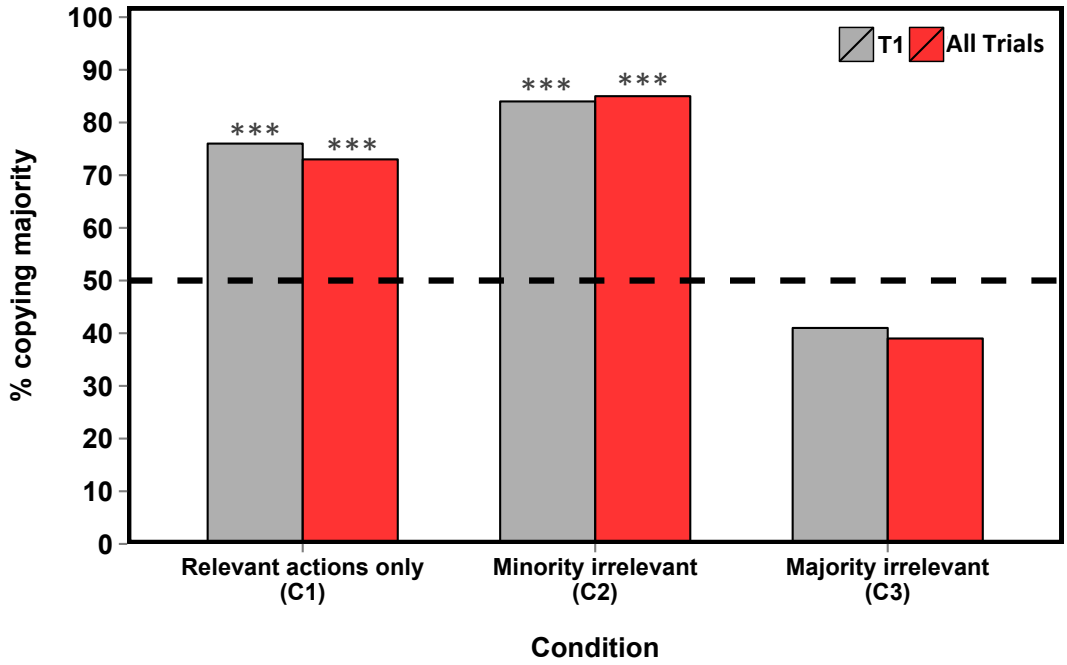


Figure 3.2. *Percentage of participants copying the majority behaviour at T1 and across all three trials (collapsed across age groups, C1- C3)*

3.4.3. Demonstrator unanimity and copying persistence across trials

Previous research suggests that children persist with performing an irrelevant action at high levels when it has been demonstrated unanimously, despite hands-on experience of task mechanics (Lyons et al., 2007; Wood et al., 2012). Children also persist in performing a unanimously demonstrated relevant solution, even when there are other equally efficacious solutions discoverable (Wood et al., 2013). Here I examined the effects of demonstrator unanimity on children's persistence across T1-T3 with the demonstrated method.

Preliminary investigations revealed age and sex effects in analyses pertaining to children's copying of irrelevant actions (reported below), but all other age and sex effects were non-significant.

Unanimous demonstrators

Within the *all irrelevant* condition (C4), where both causally relevant and irrelevant actions were demonstrated unanimously, the level of irrelevant action copying at T1 (86%) did not differ significantly from the high level of causally relevant action copying (96%; McNemar Test: $\chi^2(1) = 1.78$, $p = .18$). However, across trials (T1-T3) children were less likely to persist with the irrelevant action (70%) than the relevant action (92%; McNemar Test: $\chi^2(1) = 5.88$, $p = .02$), suggesting that fidelity erodes more quickly for irrelevant actions.

Causally relevant actions

To examine the effect of demonstrator unanimity on whether children copied causally relevant actions, I compared the behaviour of children who witnessed a unanimous demonstration of the causally relevant action (i.e., sweep *or* drawer retrieval, C2-C4 combined, N=150) with that of children who witnessed a less-than-unanimous majority (*causal actions only* condition, N=51). Children were significantly more likely to adopt the relevant action at T1 when it was unanimously demonstrated than when it was demonstrated by a less-than-unanimous majority (Unanimous=96%, Not Unanimous= 76%: $\chi^2(1) = 8.32$, $p < .004$), and were also more likely to persist with copying the unanimous demonstration across T1-T3 (Unanimous=89%, Not Unanimous= 63%: $\chi^2(1) = 16.91$, $p < .001$). Thus, children were more likely to both adopt and persist with the majority action when the demonstration was unanimous. In

other words, exposure to the alternate action, even when performed by a minority, leads to an increase in children incorporating both sweep and drawer solutions into their behavioural repertoire.

Causally irrelevant actions

Logistic regression models, including participants' sex and age, were used to examine the effect of demonstrator unanimity on children's initial copying of the irrelevant action in T1 (C2-C5, N=201), and their persistence in copying the irrelevant action across all trials (T1-T3) (C2-C4, N=150; see Figure 3.3 and Table 3.3). The *baseline* condition was included in analyses of T1 but was dropped from analyses of their persistence in copying, as it lacked the variation required to fit a logistic regression (i.e., no children in the *baseline* condition who performed the irrelevant action in T1 persisted in performing it in T2 or T3).

The frequency of demonstrators who performed the irrelevant action strongly influenced both children's initial copying of the irrelevant action in T1 (likelihood-ratio test: $\chi^2(3) = 81.20$, $p<.001$), and their persistence in copying it across T1-T3 (likelihood-ratio test: $\chi^2(2) = 51.19$, $p<.001$). Pairwise comparisons between conditions (see Figure 3.3 and Table 3.3) revealed that the high levels of initial and persistent irrelevant action copying that occurred when demonstrators were unanimous in performing the irrelevant action (*all irrelevant*: T1: 86%, T1-T3: 70%) decreased sharply in all conditions in which the demonstrators were not unanimous, including when the demonstrators performing the irrelevant action constituted a less-than-unanimous majority (*majority irrelevant:* T1: 41%, T1-T3: 21%). There was a further sharp reduction in children's initial (T1) copying of the irrelevant action when the number of demonstrators performing the irrelevant action dropped from three

(*majority irrelevant*: 41%) to just one out of four (*minority irrelevant*: 14%). In the latter case, the percentage of children who performed this action at T1 was no different from that observed in the baseline children (16%). Children who had witnessed three of four demonstrators performing the irrelevant action (*majority irrelevant*) were not significantly more likely to persist in performing it than children who saw only one of four demonstrators performing it (*minority irrelevant*), however. Thus, when the demonstrators were not unanimous, children were influenced by the number of demonstrators who performed the irrelevant action at T1, but this did not translate into differences in persisting with the irrelevant behaviour.

Table 3.3. *The effects of experimental condition and age on whether the irrelevant action was performed at T1 (C2 - C5), and persistently across T1-T3 (C2-C4)*

Model parameters	Pairwise comparisons	Estimate (S.E.)	Odds ratio
Model T1			
Intercept		0.66(1.21)^{NS}	
	All (C4) – Majority (C3)	2.25(0.51)***	9.49
	All (C4) – Minority (C2)	3.81(0.60)***	45.15
Condition[a]	*All (C4) – Baseline (C5)*	3.81(0.60)***	45.15
	Baseline (C5) – Minority (C2)	-0.005(0.58)^{NS}	1.00
	Majority (C3) – *Minority (C2)*	1.56(0.51)*	4.76
	Majority (C3) – Baseline (C5)	1.56(0.51)*	4.76
Participant's age[b]		-0.04(0.02)*	0.96
Participant's sex[c]		-0.33(0.37)^{NS}	0.72
Total model:	$R^2 = 0.46$ (Nagelkerke), $\mathcal{X}^2(5) = 84.41$, $p<.001$		
Model T1-T3			
Intercept			
Condition[a]	*All (C4) – Minority (C2)*	3.48(0.63)***	32.57
	All (C4) – Majority (C3	2.18(0.48)***	8.87
	Majority (C3) – *Minority (C2)*	1.30(0.63)^{NS}	3.67
Participant's age[b]		-0.02(0.02)^{NS}	0.98
Participant's sex[c]		-1.01(0.44)*	0.37
Total model:	$R^2 = 0.43$ (Nagelkerke), $\mathcal{X}^2(4) = 55.76$, $p<.001$		

Logistic regression models. Pairwise comparisons were undertaken with Tukey correction for familywise error.

[a] Categorical variable: (*Minority irrelevant* (C2) = only one of four demonstrators performs irrelevant action; *Majority irrelevant* (C3)= three of four demonstrators perform irrelevant action; *All irrelevant* (C4) = all four demonstrators perform irrelevant action; *Baseline* (C5) = non-social control)

[b] Numeric variable (age in months)

[c] Dichotomous variable (0 = female, 1 = male)

^{NS} p>.05; * p<.05; ***p<.001

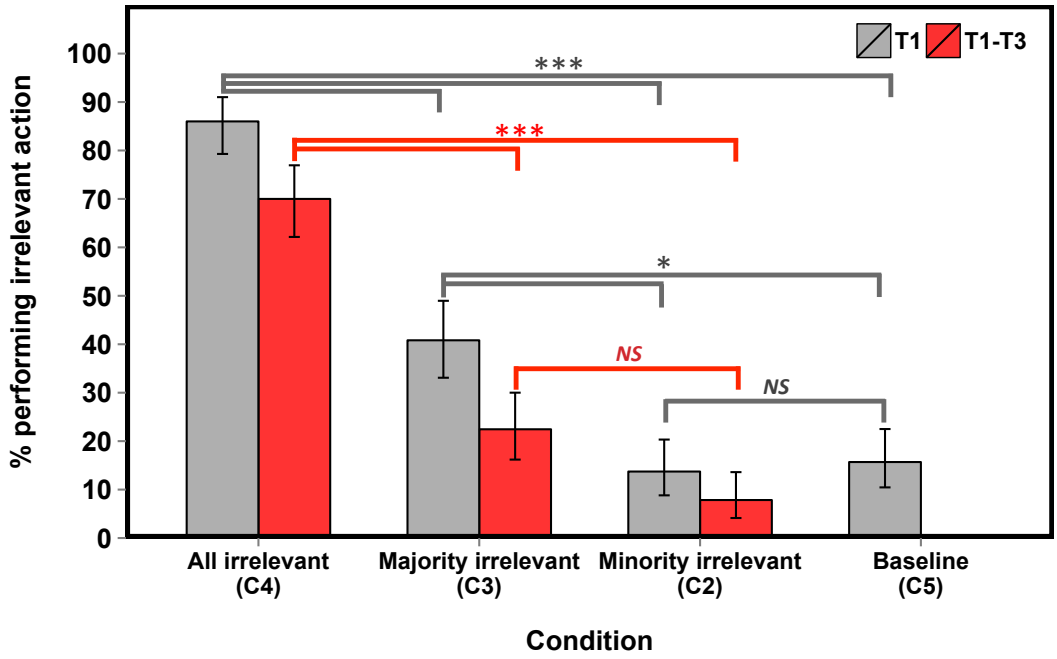


Figure 3.3. *Percentage of participants performing the irrelevant action at T1 and persistently across T1-T3 (collapsed across age groups, C2-C5)*
****p*<.001, **p*<.05, $^{n.s}$ *p*>.05. Comparisons with baseline were made at T1 only. Binomial standard errors were calculated using the Pearson-Klopper method.

Across conditions (C2-C5) children's age was found to correlate negatively with performance of the irrelevant action at T1 (see Figure 3.4 and Table 3.3), such that older children produced fewer irrelevant actions. An examination of Figure 3.4 reveals that the negative effect of age on irrelevant action *copying* (in conditions C2-C4) was confined to conditions in which the irrelevant action was not unanimously demonstrated (i.e., the *majority irrelevant* and *minority irrelevant* conditions), and was still significant following removal of the *all irrelevant* and *baseline* conditions from the analysis (C2-C3, Z = -2.04, Odds ratio = 0.95, p=.041, N=100). By contrast, children's age had no significant effect on persistence in copying the irrelevant action across T1-T3, even following exclusion of the *all irrelevant* condition. Thus the

initial (T1) tendency for increased copying of the efficient solution in older children was not maintained across repeated trials.

Although there was no effect of sex on children's initial performance of the irrelevant action in T1, boys were less likely to persist with the irrelevant action across T1-T3 than girls. Follow-up analysis revealed no interaction effect between sex and age.

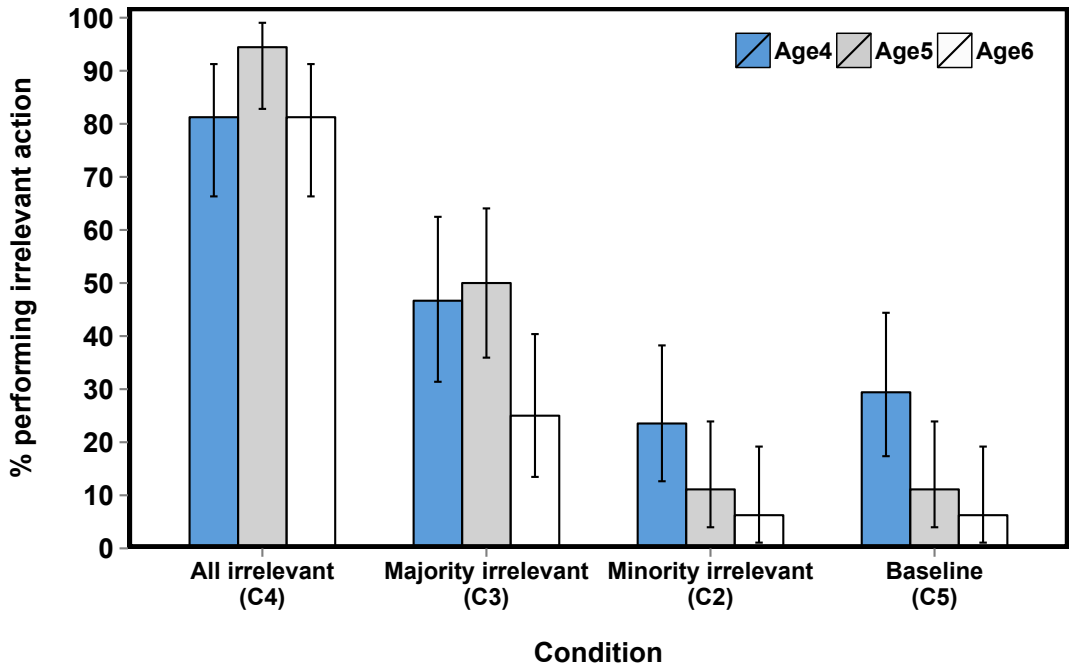


Figure 3.4. *Percentage of participants performing the irrelevant action at T1 by age group (C2 - C5)*. Age represented in years for ease of graphical representation. Binomial standard errors calculated using the Pearson-Klopper method.

3.5. Discussion

The results presented here enhance previous research findings in relation to both 'over-imitation' and majority-biased copying. They support previous results that children imitate flexibly and selectively, and provide the first direct evidence that behaviour containing causally irrelevant actions substantially alters the operation of adaptive learning biases. Within the instrumental framing of the task, majority-biased copying was always detected when the majority performed an efficient task solution, but did not extend to majorities who performed irrelevant actions. That is, majority-biased copying does not apply, at least in this context, to causally irrelevant actions.

Irrelevant action copying

In line with previous research (e.g., Lyons 2007; 2011; Horner & Whiten, 2005; McGuigan et al., 2007) children copied the irrelevant action at high levels when it was demonstrated unanimously, despite the instrumental framing of the task: "all you have to do is to try to get the sticker back out." My experimental design does not directly test between competing hypotheses concerning why children copy irrelevant information at such high levels in this context, but it does offer some insights into this.

For instance, it is unlikely that children in the *all irrelevant* condition saw the irrelevant action as causally necessary (i.e., Lyons et al. 2007; 2011), as despite copying the relevant and irrelevant actions at similar high levels initially, children were more likely to retain the causally relevant than the causally irrelevant action. Explanations based solely on assumptions about causality also imply that once the redundancy of the irrelevant action has been demonstrated, it should not show sensitivity to the relative frequency of demonstrators performing or omitting the

irrelevant actions. However, demonstrator frequency did influence children's irrelevant action copying in our study: children were more likely to perform the irrelevant action in the *majority irrelevant* than *minority irrelevant* condition. The low level of irrelevant action production in the *baseline* condition further implies that causal understanding of what was and was not required to extract the reward was not problematic for participants in any of the age groups. Considered together, these findings suggest that children's copying was influenced not by causal inferences but by demonstrator behaviour.

Older children (age 6) were less likely to copy irrelevant actions at T1 than younger children (age 4), but only in experimental conditions where irrelevant actions were not demonstrated unanimously. Previous studies in which the irrelevant action was demonstrated unanimously have found that irrelevant action copying increases with age (McGuigan et al., 2011, 2007; Nielsen & Tomaselli, 2010). A plausible explanation for these combined findings is that unanimous demonstrations generate normative pressures and susceptibilities to copy the 'way it's being done', despite the child's knowledge that it is causally unnecessary, an understanding of which increases with age (Moraru et al., 2016). However, when demonstrators vary in performance of the irrelevant action, as in our study, the pressure to conform is substantially reduced and becomes increasingly undermined by age-related increases in discarding the majority behaviour for more accurate or reliable behaviour (Einav, 2014; Seston & Kelemen, 2014).

Majority-biased copying

These results provide strong evidence that while young children do use majority behaviour as a heuristic to guide instrumental learning, they are able to do so flexibly,

calibrating their decision-making according to additional cues, such as the majority's perceived efficiency. Thus, children do not blindly follow the crowd. Majority-biased copying (regarded as a strategy for acquiring safe and effective behaviour; Boyd & Richerson, 1985, Hastie & Kameda, 2005; Henrich & Boyd, 1998) was strongest when the majority demonstrated the inefficiency of the minority's irrelevant action, and did not extend to a majority that performed irrelevant actions.

There were no significant age effects on the rate at which children copied the majority. This result contrasts with the findings of Morgan et al. (2015b), who demonstrated a positive relationship between age and majority-biased decision-making in 3- to 7-year-old children presented with a forced-choice, numerical discrimination task. However, Morgan and colleagues study did not include irrelevant actions, suggesting that age-related inferences about what is causally relevant could trump the tendency to conform in this study. These findings also contrast with McGuigan and Robertson's (2015) conclusion that children of a similar age to this study do copy irrelevant actions at high levels when demonstrated by a majority. A number of procedural differences between the two experiments could easily explain this difference; notably, the serial reinforcement of irrelevant action performance across trials in McGuigan and Robertson's design would have created increased social pressures to conform that were absent in this study.

It is possible that had the irrelevant action in the current study been demonstrated using ritualistic or normative contextual cues (e.g., Fusaro & Harris, 2008; Herrmann et al., 2013; Legg & Legare, 2015), or had the demonstrators been physically present (rather than on video), increasing children's social motivations to imitate and providing the irrelevant actions with clear socially functional properties (Nielsen, 2008; Nielsen & Blank, 2011; Over & Carpenter, 2012), a different pattern

of results, and possibly majority-biased causally irrelevant action copying, might have emerged. Future research is required to establish just where the boundaries lie in majority-biased copying.

Implications for cultural evolution

Cultural evolutionary theory states that a behavioural trait must be copied at levels proportional to the trait in the population if the trait is to be maintained at its current levels (Boyd & Richerson, 1985). These data therefore suggest that majority-biased copying could potentially stabilize functionally relevant behaviours within a population over time, but not behaviours that contain functionally redundant information. Across all three trials combined, participants who witnessed the majority perform an irrelevant action showed a bias for the minority's more efficient behaviour that approached significance. There was also no difference in persisting with copying the irrelevant action between participants who saw it performed by a non-unanimous majority or performed by a minority. In combination with the strong bias for a majority who demonstrated greater behavioural efficiency over a minority, these data imply that without additional reinforcement of the irrelevant action (e.g., sanctions, punishments, explicit teaching or other normative or social pressures), majority behaviour containing functionally redundant information will rapidly switch to a more efficient solution (i.e., irrelevant action omission), which would likely continue to increase towards fixation.

It is plausible that had the causally irrelevant action in this study encompassed more substantial efficiency costs, as might be true of many naturalistic behaviours, lower rates of irrelevant action copying and faster rates of erosion over time would have been observed; a suggestion consistent with the findings of Keupp et al. (2016).

Varying the ratio of majority versus minority demonstrators who performed the irrelevant action (for example 25:1 instead of 3:1), would also plausibly affect the rate of erosion, as would manipulating the relative age (Wood et al. 2012), status (McGuigan, 2013), or group membership (Oostenbroek & Over, 2015), of the demonstrators. Examining the interaction of different types of learning biases in irrelevant action copying is also an area ripe for future research.

In sum, to my knowledge, this chapter presents the first evidence that young children flexibly and adaptively adopt a majority-biased learning strategy when faced with an instrumental learning goal and the opportunity to integrate social information from multiple individuals. Majority-biased copying did not extend to causally inefficient and irrelevant actions, despite these being copied at high levels when demonstrated unanimously. When just one individual dissented from the majority, 'over-imitation' plummeted. Thus, these data suggest that the presence of causally irrelevant actions might substantially alter the operation of adaptive learning biases. This finding has obvious implications for cultural evolutionary theory; namely that, in more real-world contexts in which actions are not always demonstrated unanimously, causally irrelevant, and potentially costly, actions are unlikely to be maintained in causal or instrumental contexts.

But instances of copying causally irrelevant actions, to serve social, ritualistic, or normative purposes, clearly do exist. These data also provide support for the operation of socially driven motivations in causally irrelevant action copying, despite the instrumental framing of the task. However, I suggest the term 'over-imitation' is inaccurate and misleading when causally irrelevant actions encompass social or

normative functional properties, as their performance in this instance no longer represents puzzling or irrational behaviour.

To the contrary, these findings illustrate a flexible and highly functional integration of social learning strategies, through which individuals combine social and non-social sources of information to home in rapidly on the relevant actions in instrumental tasks, while remaining sensitive to the social functions of imitation. This suggests that our species might more accurately be cast as 'optimal' rather than 'over'-imitators.

Chapter 4

Cooperation and Imitation: A Series of Experimental Investigations in Adults

Abstract

The previous chapters considered questions pertaining to the role of high-fidelity learning mechanisms (e.g., imitation) in the transmission and generation of cultural behaviour. In this chapter, the focus shifts to the relationship between cooperation and imitation, and its role in generating group cohesion and cultural behaviour. In humans it is hypothesised that simple (automatic) imitation and cooperation are engaged in a mutually reinforcing virtuous circle, which might act as a springboard for the development of more complex forms of cultural learning and cooperation (e.g., Heyes 2012a). The pathway leading from imitation to increased cooperation is well documented, while the reciprocal relationship, from cooperation to increased imitation, is little studied. Using a series of three experiments, I attempted to test the hypothesis that increases in cooperative rapport lead to increases in imitative behaviour between interaction partners. In the first experiment, imitative behaviour was positively associated with a measure of the quality of participants' cooperative interaction (cooperative task performance), but was not higher overall in participants who cooperated relative to a control group (who performed the task alone). Moreover, participants in the control group imitated their interaction partner more if they had performed poorly; consistent with the hypothesis that poor individual performance leads to increasing desire to affiliate with others through imitation. However, these effects were not replicated in Experiments 2 and 3, and nor was imitative behaviour higher overall in participants who experienced cooperation, despite a series of incremental modifications designed to enhance participants' cooperative experience. Future studies are required to examine and identify possible nuances in the relationships tested here.

4.1. Introduction

It is well established that people subconsciously imitate each other's mannerisms, gestures and expressions. This phenomenon – known as the chameleon effect (Chartrand & Bargh 1999), non-conscious mimicry (van Baaren et al., 2009), response facilitation (Byrne, 2002), or automatic imitation (Heyes, 2011) – is believed to result in enhanced social interactions. For instance, when experimental participants are imitated by a confederate, the participants report that they like the confederate more than do participants who were not imitated (Chartrand & Bargh 1999). Imitated participants also report that they found the interaction more enjoyable (Tanner et al., 2008), and rate their interaction partner as more persuasive (Van Swol, 2003), than participants who were not imitated.

The benefits of being imitated appear to not just advantage the imitator. In a negotiation study by Maddux et al. (2008), dyads in which one subject was instructed to mimic their partner's mannerisms secured better outcomes for both partners, compared with dyads that had no imitator. Participants who have been imitated are also more likely than non-imitated participants to help other group members with simple tasks (e.g., picking up objects another has dropped) – helping both the person who has imitated them or a third party – and make increased anonymous donations to a charitable cause (van Baaren et al. 2004). Even children as young as eighteen months old will help others more (e.g., pick up something dropped) when they have been imitated than children who have not been imitated (Carpenter et al., 2013).

Thus, even when it is deliberately and experimentally orchestrated, the pro-social effects of imitation appear pervasive. The relationship between imitation and cooperation is unlikely to be unidirectional, however: While imitating another leads to

increased cooperation, being cooperative is also thought to increase imitation, thus engaging automatic imitation and cooperation in a 'virtuous circle' of mutual reinforcement (Heyes, 2012a). This virtuous circle, in which persons are neither aware of being imitated nor of imitating (Chartrand & Bargh, 1999), might function to help maintain large-scale collective action and information exchange, by acting as a social glue between members of a cultural group. Experimental evidence that individuals are less likely to imitate members of other ethnic (van Baaren et al., 2009) or religious (Yabar et al., 2006) groups than members of their ingroup, lends some support to this suggestion.

Although the mechanisms that give rise to the virtuous circle are still poorly understood, it is plausible that the processes responsible for engaging simple imitation and cooperation in this reinforcing loop have acted as a springboard for the coevolution of more complex forms of cultural behaviour and cooperation (Heyes, 2012a). Indeed, experimental studies suggest that simple, automatic imitation operates jointly with more complex forms of rational and intentional imitation in subjects engaged in an instrumental gesture making game (Belot et al., 2013). It is then possible that an escalating cycle of automatic imitation and cooperativeness between interaction partners could bootstrap the emergence of more complex forms of social learning (including rational and intentional imitation, and teaching through language; Byrne & Russon, 1998; Shea, 2009; Tomasello, 1999) along with ever-increasing forms of ultra-cooperativeness.

However, despite the proposed mutual reinforcement between cooperation and imitation, the pathway leading from cooperation to imitation is less documented and established than its reciprocate. Perhaps the most convincing evidence was shown by Leighton et al. (2010), who found that individuals primed with words such as

"cooperate" and "together" showed higher levels of automatic imitation than those primed with words such as "alone" and "selfish" (using a stimulus-response compatibility procedure that required them to produce matching or mismatching hand movements in response to a stimuli). People also imitate a person who they like more than a person they don't like *a priori* (Likowski et al., 2008; McIntosh, 2006; Stel et al., 2010), and incur reputational costs when a third party observes them imitating an inhospitable other, relative to imitating a cordial other (Kavanagh et al., 2011). However, participants do not preferentially imitate an 'agreeing' person more than a 'disagreeing' person, when faced with the two simultaneously (Van Swol, 2003), and also increase their imitative tendencies towards groups from whom they have been excluded, rather than included (Lakin et al., 2008; see also Watson-Jones et al., 2015). Thus, the social functions served by imitation in forging and repairing cooperative relationships (Lakin & Chartrand, 2003; Over & Carpenter, 2012) might sometimes offset the immediate positive relationship expected between these two traits.

The purpose of the experiments in the current chapter was to examine the effect of cooperative interactions on automatic imitation in more detail. The work in this chapter is presented as a series of three experiments that reflect the chronological development and aims of the work undertaken. Specifically, in Experiment 1 and Experiment 2, I compared the imitative behaviour of participants who performed a coordination task cooperatively with a confederate, with participants in a control condition who performed the coordination task alone. In Experiment 3, pairs of naïve participants performed the coordination task with each other (or alone), rather than with a confederate. The first two experiments employed a naturalistic measure of imitation (i.e., face touching; e.g., Chartrand & Bargh, 1999), while a gesture-matching measure was adopted from Belot et al. (2013) in Experiment 3. In

manipulating participants' engagement in an actual collaborative activity, this study extends earlier work that examined the effect of experimentally priming a cooperative mind-set on participants' propensity to imitate images of gestures presented via a computer display screen (Leighton et al., 2010).

Across all experiments, I hypothesised that participants who undertook the coordination task cooperatively with an interaction partner would show a greater propensity to imitate the interaction partner's behaviour, than participants who had not previously been given the opportunity to form a cooperative rapport (Chartrand & Van Baaren, 2009; Heyes, 2012a; van Baaren et al., 2009). However, as the coordination task presented the possibility for variable performance between participants (apparatus, see methods), I further anticipated that participants' propensity to imitate their interaction partner might show sensitivity to their performance on the task. I expected that performance would vary positively with imitative behaviour when participants had performed the task cooperatively, reflecting the quality of their cooperative interaction. Conversely, I anticipated that participants who performed the task alone would not display the same positive relationship between task performance and imitative behaviour, and instead might show the reverse pattern (i.e., a negative relationship between performance and imitative behaviour) if poor individual performance led to an increase in desire to affiliate and build rapport with others (Chartrand & Van Baaren, 2009).

4.2. Experiment 1

In the first experiment, participants first undertook the coordination exercise (the buzz wire, see apparatus below), either with a confederate or individually, and then

engaged in a separate communications exercise with the confederate. During the communications exercise a naturalistic measure of subconscious imitation was recorded (face touching).

4.2.1. Methods

4.2.1.1. Participants

Fifty-four participants were recruited from the St Andrews University population via the Psychology Department's online sign-up system SONA. An additional 11 participants took part, but were excluded due to: noticing the concealed camera (2); not paying attention during the communication exercise (1); suspecting that the other 'participant' was a confederate (1); resting their face on their hands throughout the face touching phase (1); and noticing the confederate's face touching mannerisms (6). Participants had a mean age of 22 years and 14 (26%) were male. There were no differences in participants' sex (Fisher's Exact Test: p=0.76), or age (*Welch's* $t_{(37.95)}$=1.54, p=0.13) between conditions. All subjects were awarded a participation fee of £5, which they could increase up to £10 for good performance in the first phase of the experiment. Ethical approval for this study was given by UTREC of the University of St Andrews.

4.2.1.2. Apparatus (coordination task)

The buzz wire task, a coordination exercise, was designed and custom-made for the purpose of this study (Figure 4.1). The buzz wire could be performed jointly (Figure 4.1a) or individually (Figure 4.1b), and the goal was to move a metal wand, which was attached to the wire by a small metal ring, backwards and forwards along the

buzz wire, from one end to the other, without touching it. Upon touching the wire, a buzzer sounded, at which point participants were required to return the wand to the starting end of the wire and begin again. A successful run of the buzz wire was defined as moving the wand from one end of the wire to the other without sounding the buzzer. This novel task was thought particularly suitable for the purposes of this study as, when performed jointly, the motor requirements of the two roles are essentially identical and therefore encourage synchronous, coordinated and cooperative behaviour.

Figure 4.1. *Participant undertakes the buzz wire task (a) jointly with another person, or (b) alone*

4.2.1.3. Design

In a between-subject design, participants were randomly allocated to the *cooperation* or *control* condition. The experiment consisted of two stages: the first phase provided the experimental manipulation (i.e., *cooperation* or *control*), while the second stage,

which was identical across conditions, provided the measure of behavioural imitation. Data collection was balanced in relation to two females who acted as confederates, who were unknown to participants, and who were blind to the hypotheses of the study.

In the first stage, participants in the *cooperation* condition undertook the buzz wire task cooperatively with a confederate (see Figure 4.1a), while participants in the *control* condition undertook the buzz wire task alone (Figure 4.1b). Methods for the second stage of the experiment were adapted from previous experimental protocols that have been used to study non-conscious imitation of face touching behaviour (Chartrand & Bargh, 1999; Lakin & Chartrand, 2003; Yabar et al., 2006). Participants were video recorded while they watched a "live feed" of the confederate communicating information from an adjoining room (actually recorded earlier), during which the confederate was touching her face. The same confederate video footage was used throughout the experiment to ensure consistency within and between experimental conditions, and confederates were also issued with scripts to ensure that they each communicated identical information.

4.2.1.4. Procedure

On arrival each participant was informed that they would be taking part in two unrelated experiments with another participant (the confederate). In the first stage, participants were informed that they would undertake a coordination exercise. Those assigned to the *cooperation* condition were required to work jointly with the confederate in completing the buzz wire task. During this phase, which lasted for 10 minutes, participants were permitted to communicate verbally about the task, and encouraged to achieve as many successful runs of the buzz wire as possible in the

allotted time. Participants in the control condition completed the same buzz wire task alone, and in a separate room from the confederate. All participants were informed that their final pay-off was directly related to their performance in this phase, and that they would receive an extra £0.25 for every successful run of the buzz wire, in addition to the £5 participation fee.

Before moving on to the second stage of the experiment, the experimenter informed participants that she had forgotten to complete some paperwork (a short questionnaire regarding whether they had any previous experience on a coordination task such as the buzz wire) that required their (dated) signature. In an adaptation of a previous method used to study conformity (Coultas, 2004), the confederate first signed and dated the form using an analogical method of date signing (e.g., 14th March 2012), before passing it on to the participant to complete. Coultas (2004) previously found that experimental participants switched from the more common practice of writing the date numerically (14/03/2012), to signing analogically, when they believed others were signing analogically. The method was adapted here to serve as an additional measure of the participant's willingness to imitate the confederate. The consent form that participants signed and dated at the start of the experiment served as a measure of their baseline method of date signing.

In the second stage, participants were instructed that they were partaking in a communications exercise, during which they would watch a "live feed" of the other participant (actually the confederate videoed earlier), communicating information about four photographs from another room (the participant's goal was to later identify the photographs described from a larger series of photographs). Each of the four photograph scripts communicated by the confederate described a landscape, and included details of persons, but no emotional expressivity (e.g., a group of people

hiking). The first 2 minutes of video playback footage were used to ascertain a baseline measure of each participant's face touching behaviour. During the baseline phase, the confederate examined the pictures in preparation for the description task, but she did not touch her face. Immediately following the baseline period, the confederate began touching her face during and between verbal descriptions of the pictures: a period lasting approximately 4.5 minutes (imitation phase). Participants were surreptitiously filmed throughout the video playback, which they watched from a chair placed 50 cm away from a 20" LED display screen.

Next, participants completed a debrief questionnaire probing for (a) general suspicions about the confederate or procedure (b) what they thought the experiment was about, and (c) whether they noticed any particular mannerisms exhibited by the confederate. They were then thoroughly debriefed on the purpose of the experiment, and given the opportunity to ask questions. The experimental procedure lasted approximately 40 minutes. Participants were then free to collect their payment and leave.

4.2.1.5. Coding and analysis

The experimenter live-coded the participants' performance on the buzz wire task (i.e., the number of successful wire runs), and coded 100% of participants' face touching behaviour from video footage. Following previous protocols (Chartrand & Bargh, 1999; Lakin & Chartrand, 2003) the amount of time (in seconds) that each participant spent touching their face during the baseline and imitation phase of the video playback was coded. An independent coder blind to condition and hypotheses scored 20% of videos, and inter-observer agreement was high: Chronbach's alpha= 0.87.

All analyses were carried out in R version 3.1.3. Quasi-binomial models were used to examine the proportion of time that participants engaged in face touching during the experimental phase. A quasi-binomial model, which employs an additional parameter to adjust for deviations from the variance structure of the binomial distribution (Crawley, 2013), was preferred, as a conventional binomial model displayed evidence of overdispersion during model fitting. Quasi-AIC (QAIC) values were obtained using the package *MuMIn* (Barton, 2015), and a reduction in quasi-AIC of more than 2 was regarded as evidence that inclusion of an independent variable significantly improved model fit. Zero-inflated negative-binomial models were fitted with the R package *pscl* (Zeileis et al., 2015).

4.2.2. Results

Participants in the *cooperation* condition produced higher scores on the buzz wire task (mean=9.94) than those who were in the *control* condition (mean=5.21, *Welch* $t_{(51.8)}$=5.27, p<0.001, N=54). Thus performing the buzz wire task jointly with another person was beneficial for performance. Participants who performed the buzz wire task jointly with the confederate did not afterwards copy the confederate's analogical method of date signing more frequently than participants who undertook the buzz wire task alone (Fisher's Exact Test: p=0.28, N=48; 8 participants who had signed the date analogically in the baseline were excluded from this analysis).

To examine the effects of experimental condition and buzz wire performance on face touching imitation, a quasi-binomial model was run on the proportion of time the participant engaged in face touching during the imitation phase. A covariate to adjust for each participant's baseline measure of proportional time spent face touching

was also included. Other variables included in the model were condition (cooperation/control), buzz wire performance (the number of successful runs), and the interaction between condition and buzz wire performance. Although the dependent variable exhibited a high proportion of zeros (26% of all participant responses), a zero-inflated negative-binomial model did not yield results that differed in direction or significance from those reported for the quasi-binomial model.

Figure 4.2 demonstrates that, after controlling for the significant effect of participants' baseline level of face touching (t=7.08, odds ratio= 2.10, p<.001, N=54), there was a significant interaction effect between buzz wire performance and condition, suggestive that face touching imitation increased with increasing performance when participants had undertaken the task cooperatively, but decreased with increasing performance when they had performed the buzz wire task alone (Δ_{QAIC}= -2.34, t=-2.07, $p<.04$; see Figure 4.2b). There was no main effect of experimental condition following removal of the significant interaction term (t=1.16, odds ratio= 1.85, p=.25).

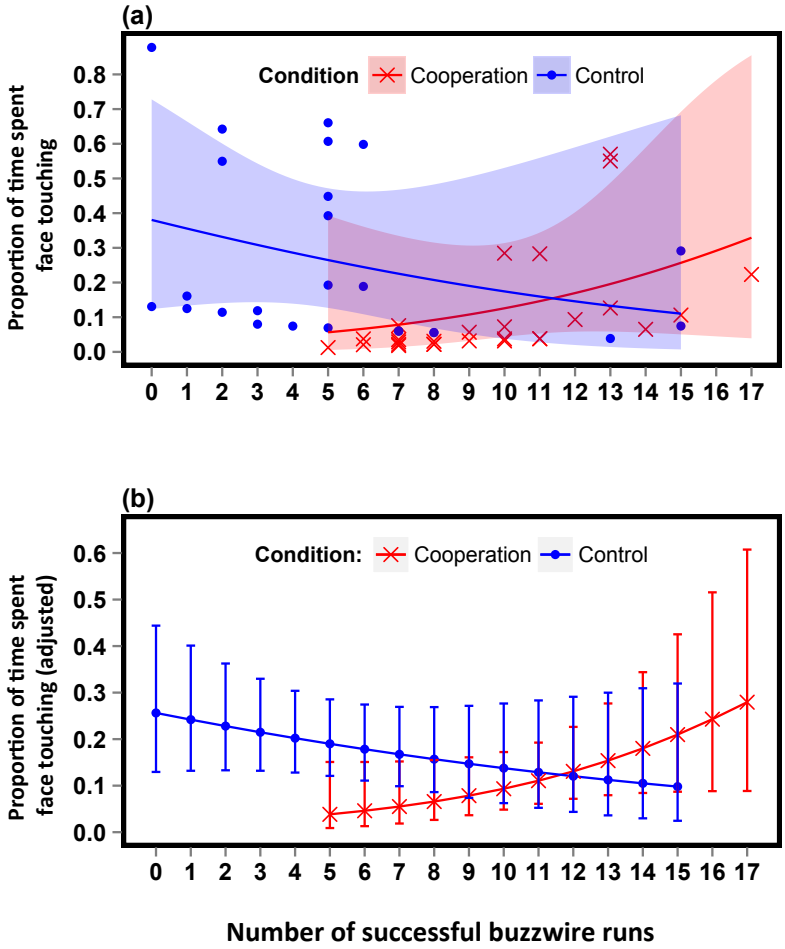


Figure. 4.2. *(a) Scatterplot showing the proportion of time spent engaging in face touching behaviour in the imitation phase by condition and performance on the buzz wire task (binomial lines of best fit ±95% CIs; data not adjusted for participants' baseline level of face touching); and (b) predicted mean values ±95% CIs (adjusted for participants' baseline level of face touching)*

4.2.3. Discussion

Experiment 1 showed evidence of an interaction effect between buzz wire performance and automatic imitation in the direction predicted: When participants had undertaken the buzz wire task cooperatively, imitation of the interaction partner (confederate) was positively associated with performance, but negatively associated with performance when participants had undertaken the buzz wire task alone. This

relationship was expected if task performance in the *cooperation* condition was positively related to the quality of cooperative rapport established, while poor task performance in the *control* condition led to an increase in participants' desire to affiliate with and imitate others. However, contrary to expectations, automatic imitation was not higher overall in participants who had been given a prior opportunity to establish cooperative rapport with the confederate.

A possible explanation for this discrepancy between results and predictions might relate to participants' experience of the task. While it was initially envisaged that the cooperation task used in this set-up would encourage participants to develop a feeling of collaboration and behavioural synchronization with the confederate, it became apparent during Experiment 1 that some participants in the *cooperation* condition found working with the confederate stressful (occasionally mentioning concerns that the confederate was hindering their performance on the buzz wire task or vice versa). Participants in the *cooperation* condition were also unaware of the performance advantages bestowed by cooperation versus performing the buzz wire task alone (cooperation approximately doubled average performance) – an additional factor that might have prevented the expected positive association between cooperation and imitation. In Experiment 2, I attempted to address these issues.

4.3. Experiment 2

The aims and design of Experiment 2 closely followed Experiment 1, but offered a number of procedural modifications. In an attempt to enhance the relationship between cooperation and imitation, participants in Experiment 2 who performed the task cooperatively were provided with explicit information about the task-related

benefits of cooperative performance (i.e., information about the lower average scores of *control* participants), which had not been available in the first experiment. In Experiment 2, I also implemented a number of measures designed to reduce any anxiety that participants might have experienced when undertaking the buzz wire task cooperatively, and included additional questions in the end-of-study questionnaire to assess participants' experience of this. I expected the general pattern of results to be comparable with Experiment 1, but with the exception that overall levels of imitation would be higher in the *cooperation* condition than in the *control*.

4.3.1. Methods

4.3.1.1. Participants

Thirty-eight participants were recruited from the St Andrews University population via the Psychology Department's online sign-up system SONA. An additional 12 participants took part, but were excluded due to: noticing the concealed camera (3); not paying attention during the communications exercise (2); suspecting that the other 'participant' was a confederate (2); resting their face on their hands during the face touching measure (2); and noticing the confederate's face touching mannerisms (3). Participants had a mean age of 22 years and 10 (26%) were male. There were no differences in participants' sex (Fisher's Exact Test: p=0.46), or age (*Welch's* $t_{(19.5)}$=1.01, p=0.32) between conditions. Ethical approval for this study was given by UTREC of the University of St Andrews.

4.3.1.2. Design and procedure

The design and procedure of Experiment 2 followed that of Experiment 1, with the following exceptions. Participants in the *cooperation* condition were informed of the average score of participants who had performed the task individually (based on data from the *control* condition in Experiment 1), in order to highlight the performance-related advantages of performing the buzz wire task cooperatively. Participants in the *control* condition were given no information pertaining to performance.

Additionally, in an attempt to reduce any stress or anxiety associated with performing the buzz wire task, a practice period of 4 minutes was introduced for all participants in Experiment 2, which did not contribute to their performance-related payoff. During the practice period, participants were free to try adjusting the height of the buzz wire, along with different techniques of moving the wand, while the experimenter busied herself with another task. Participants in the *cooperation* condition performed the practice period jointly with the confederate, who was instructed to explicitly consult and cooperate with the participant on their preference for the height of the buzz wire and wand movement.

After the practice period had ended, participants then spent a further 6 minutes performing the buzz wire task, and were aware that they would receive an extra £0.50 for every successful run of the buzz wire during this period. In the *cooperation* condition, on reaching the average score of a participant assigned to the control condition, the confederate announced, "at least we're doing as well as somebody on their own." This was followed by, "now we're doing better than an individual person," when the average score was exceeded.

The procedure for the second stage (i.e., face touching imitation) was identical to Experiment 1. In addition to the previous questionnaire (i.e., questions probing for

general suspicions about the experimental procedure), participants were asked to answer the following three extra questions: (1) How much did you like the other person? (2) How well do you think you performed on the buzz wire task? And (3), How cooperative do you think the other person was during the buzz wire task? All answers to the additional questions were provided on a 7-point likert scale.

4.3.1.3. Coding and analysis

Coding and analysis followed the same procedure as Experiment 1. The experimenter coded 100% of the video footage detailing participant's face touching behaviour. An independent coder blind to condition and hypotheses scored 20% of videos, and inter-observer agreement was high: Chronbach's alpha= 0.84.

4.3.2. Results

Participants in the *cooperation* condition produced higher scores on the buzz wire task (mean=5.47) than those who were in the *control* condition (mean=2.29, *Welch* $t_{(30.1)}$=4.72, p<0.001, N=38). Thus performing the buzz wire task jointly with another person was beneficial for performance. Participants who performed the buzz wire task jointly with the confederate did not afterwards copy the confederate's analogical method of date signing more frequently than participants who undertook the buzz wire task alone (Fisher's Exact Test: p=1, N=30; 8 participants who had signed the data analogically in the baseline were excluded from this analysis). All participants in the *cooperation* condition rated the other participant's (i.e., the confederate's) cooperativeness as six or seven on the seven point likert scale (seven indicating very cooperative).

As in the first experiment, a quasi-binomial model was run on the proportion of time the participant engaged in face touching during the imitation phase, with baseline level of face touching, along with experimental condition and buzz wire performance (including their interaction), as independent variables. In contrast to the first experiment, and after controlling for the significant effect of participants' baseline level of face touching (t=2.79, odds ratio= 1.79, p<.01, N=38), there was no interaction effect between buzz wire performance and condition (t=0.48, p=*.64*; see Figure 4.3). Moreover, Figure 4.3b indicates that the non-significant interaction effect in Experiment 2 showed an opposite pattern to that found in the first experiment. Following removal of the interaction term from the model, there was also no main effect of buzz wire performance (t=-0.53, odds ratio= 0.63, p=.60, N=38), nor of experimental condition (t=-0.20, odds ratio= 0.97, p=.85, N=38). A separate analysis of face touching behaviour, also employing quasi-binomial models but adopting the questionnaire-based measures as independent variables, revealed no significant effect of how much the participant liked the confederate (t=-0.74, odds ratio= 1.0, p=.99, N=38), nor how well the participant thought they had performed on the buzz wire task (t=-0.74, odds ratio= 0.80, p=.47, N=38), after controlling for baseline levels of face touching.

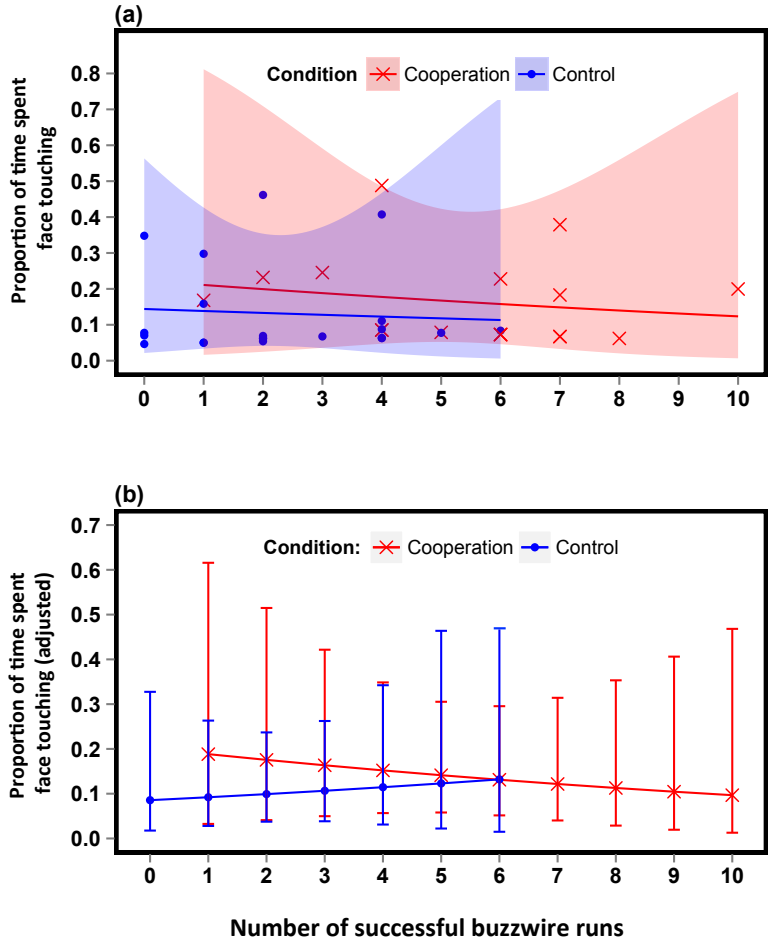


Figure. 4.3. *(a) Scatterplot showing the proportion of time spent engaging in face touching behaviour in the imitation phase by condition and performance on the buzz wire task (binomial lines of best fit ±95% CIs; data not adjusted for participants' baseline level of face touching); and (b) predicted mean values ±95% CIs (adjusted for participants' baseline level of face touching)*

4.2.3. Discussion

Experiment 2 failed to replicate any of the patterns observed in Experiment 1. Additionally, imitative behaviour in the *cooperation* condition was not significantly above that observed in the non-cooperative control, despite additional interventions (relative to Experiment 1) to reduce any anxiety caused by performing the buzz wire task cooperatively, and interventions to raise awareness of the performance-related

benefits of performing the task cooperatively. This was observed despite all participants assigned to the *cooperation* condition rating the confederate highly for cooperativeness (i.e., 6 or 7 on a 7-point likert scale). Unlike previous studies (Likowski et al., 2008; McIntosh, 2006; Stel et al., 2010), there also appeared to be no effect of how much the participant reported liking the confederate on imitative behaviour, although this might reflect differences in the way that 'liking' was assessed between studies.

The non-repeatability of general patterns across the first two experiments raises doubts as to the robustness of the interaction effect detected in Experiment 1. Given that the non-significant findings in Experiment 2 displayed an opposite trend to that observed in Experiment 1, it is unlikely that the failure to replicate the initial findings is simply an artefact of the small sample size (i.e., N=38) in Experiment 2. Experiment 1 utilised a slightly larger sample (N=54), but the interaction effect detected only just passed the significance threshold ($p<.04$), raising the possibility that this result might reflect stochastic or systematic error. It is unlikely that the procedural differences between Experiment 1 and 2 would explain the opposite pattern of results, as the modifications implemented were simply designed to increase overall cooperative rapport in the *cooperation* condition, while the *control* condition remained largely unchanged. The sample sizes employed in the current study are on par with other published studies that utilise very similar experimental designs (e.g., Chartrand & Bargh, 1999 (N=35); Lakin & Chartrand, 2003 (N=47); Yabar et al., 2006 (N=26)), and illustrate the risk of non-repeatability with small samples, but also the difficulty of generating large samples when the procedure for obtaining each data item is detailed and long.

A further potential criticism of Experiment 1 and Experiment 2 is that a forced cooperation task in experimental settings might not easily lend itself to the expected relationship between cooperation and imitation, and that a voluntary cooperation exercise, where participants are free to choose to cooperate, is more representative of naturalistic situations. Allowing participants to choose or decline cooperation is also more in line with studies that have tested the reverse relationship, from imitation to cooperation, and found that participants who have been imitated by a confederate are more likely to *voluntarily* help this person (e.g., van Baaren et al. 2004). The use of naturalistic measures in studies of automatic imitation (e.g., face touching, foot tapping, pen handling), which was adopted in Experiments 1 and 2, has also been criticised for failing to provide an accurate measure of imitation: the effect of increased face touching might more accurately reflect stimulus enhancement effects, for example, as the precise motor actions of the 'imitator' are not expressly assessed (Leighton et al., 2010).

In Experiment 3, I attempted to address these remaining concerns.

4.3. Experiment 3

In Experiment 3, pairs of naïve participants interacted with each other rather than with a confederate. Procedural and design modifications permitted active decision-making between participants regarding cooperative or individual performance on the buzz wire task (i.e., cooperation was chosen rather than enforced by the experimenter). These design modifications also presented the opportunity to use a different imitation measure, based on hand gesture matching between participants (the Matching Pennies game, see Methods below), which afforded a more reliable assessment of imitative

behaviour than is offered by naturalistic measures such as face touching (Leighton et al., 2010).

All pairs of participants in Experiment 3 were provided with the opportunity to practice performing on the buzz wire task both jointly (cooperatively) and individually, after being allocated to a condition that either encouraged or discouraged cooperation. Participants could then decide to complete the trial by performing the task cooperatively or individually. In line with the first two experiments, I expected that participants who performed the task cooperatively would show higher levels of imitation than those who chose not to cooperate. However, contrary to Experiment 1 and Experiment 2, I did not predict or test for an interaction between experimental condition and performance, as each participant's decision to cooperate or not in Experiment 3 was based on a comparison of their relative practice period experience of cooperative versus individual performance, and not restricted to an individual *or* cooperative experience of the buzz wire task, as per the first two experiments.

4.3.1. Methods

4.3.1.1. Participants

Forty participants (20 pairs) were recruited as per the same procedure as the first two experiments. Participants had a mean age of 21.9 years and 35% were male. There were no differences in age (*Welch's* $t_{(20.0)}$=0.08, p=.94) or sex (Fisher's Exact Test: p=1) between conditions. Participants were not permitted to take part in the same session with another participant whom they already knew. Ethical approval for this study was given by UTREC of the University of St Andrews.

4.3.1.2. Matching Pennies game (imitation measure)

The Matching Pennies game (adapted from Belot et al. 2013) was used to provide a measure of automatic imitation in Experiment 3. Matching Pennies is a zero-sum hand gesture game (i.e., one player's win automatically signals the other player's loss) that closely resembles the widely known game rock-paper-scissors. The Matching Pennies game is played by pairs of players, and begins with both players in the neutral START position (Figure 4.4a). After a count of three, each player must choose to execute either the CLOSED-HAND gesture (Figure 4.4b) or the OPEN-HAND gesture (Figure 4c). Players are given different incentives during gameplay: if Player 1's incentive is to produce matching gestures with Player 2, then Player 2's incentive is to produce mismatching gestures, and vice versa. The Nash equilibrium for each player in Matching Pennies is therefore to produce the OPEN- or CLOSED-HAND gestures each with a probability of 0.5. During gameplay, one of the players is sighted, while the other is blindfolded (i.e., the blindfolded player's role is simply to provide the gestures that the sighted player might imitate). Despite being instructed to produce their gestures at the same time, players sometimes display fractional discrepancies in their timing, providing an opportunity for automatic imitation to operate on the sighted player when the blindfolded player presents early (Cook et al., 2012). Belot et al. (2013) showed that players of Matching Pennies match the gesture of their opponent significantly greater than would be expected by chance, and have attributed this effect to the occurrence of automatic imitation in the sighted player. In Belot and colleagues' study, automatic imitation was detectable even when the sighted player was incentivised to mismatch gestures, although players matched more frequently when the sighted player was incentivised to match, due to the additional operation of intentional gesture imitation.

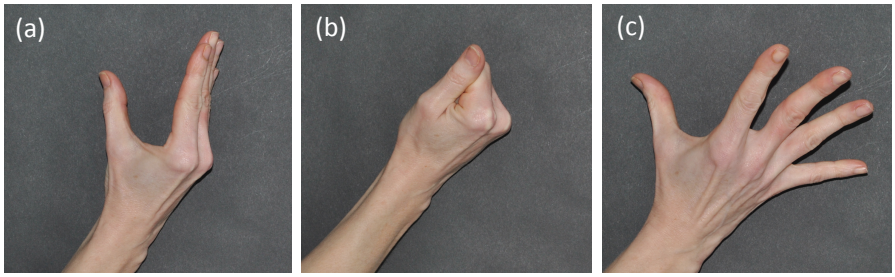


Figure 4.4. *Hand gestures used in the Matching Pennies game, detailing (a) the neural START position, (b) the CLOSED-HAND gesture, and (c) the OPEN-HAND gesture (adapted from Belot et al. 2013)*

4.3.1.3. Design

In a between-subjects design, pairs of participants were randomly allocated to the *cooperation encouraged* condition or the *cooperation discouraged* condition. The experiment took part in two stages. The first stage provided the experimental manipulation, and involved a modified version of the previous buzz wire task. The second stage, which was identical across conditions, provided the measure of behavioural imitation, and employed the Matching Pennies game (see above).

4.3.1.4. Procedure

On arrival, each participant was introduced to the other participant and informed that they would be taking part in two unrelated experiments. In the first stage, participants engaged in a four-minute practice period on the buzz wire task, before together choosing whether to complete the exercise: (i) cooperatively (joint-action wand, Figure 4.1a), or (ii) individually (individual-action wand, Figure 4.1b). Crucially, in

the *cooperation encouraged* condition (number of pairs =10; 3 pairs were mixed sex) the circular attachments of the wands were manipulated to encourage a preference for cooperative working (diameter of joint-action wand attachment: 3 cm; diameter of individual-action wand: 1.5 cm). In the *cooperation discouraged* condition (pairs=10; 4 pairs were mixed sex) participants followed the same procedure but the wands were manipulated to discourage a preference for cooperative working (diameter of joint-action wand attachment: 1.5 cm; diameter of individual-action wand: 3 cm). Participants in both conditions spent 2 minutes practicing with the individual-action wands and 2 minutes practicing with the joint-action wand, before deciding whether to work together or alone. After participants had decided, they performed the exercise cooperatively or individually for 6 minutes. All participants were informed that their final pay-off was directly related to their performance in this phase, and that they would receive an extra £0.50 for every successful run of the buzz wire, in addition to the £5 participation fee.

In the second phase of the experiment, participants played the Matching Pennies game (adapted from Belot et al., 2013, see above). The experimenter provided participants with detailed verbal instructions about the rules of the Matching Pennies game, checking that they understood and could create the required hand gestures before beginning. Within each participant pair, individuals were randomly labelled Player 1 or Player 2, and the experimenter acted as the game umpire. Each participant pair played four rounds of 20 games, and participants were sitting facing each other throughout each round. The game incentive of each player (matching vs. mismatching gestures), and the player who was blindfolded, was counterbalanced between rounds (see Table 4.3.1).

At the start of each new round, players played three practice games, during which both players were sighted. After this, only one player remained sighted (the other wore the blindfold) and game play began. In each game, players were required to produce their hand gesture at the experimenter's count of three, following which the experimenter announced the game winner (e.g., "Player 1") and recorded the result on a score sheet, before players resumed the neutral START position, ready for the next game. Rounds in which there was an obvious violation of the game rules (e.g., delayed or premature gesture execution) were replayed (a total of 8% of games were replayed). All participants were informed that they would receive an additional 5p for every Matching Pennies game that they won, thus a player who won all of their games would receive an additional 0.05 x 80 = £4.

At the end of the experiment, participants were provided with a full debrief and an opportunity for further questions. The experimental procedure lasted approximately 40 minutes. Participants were then free to collect their payment and leave.

Round	**Game Incentive**		Player Blindfolded	Num of Games
	Player 1	Player 2		
1	Match	Mismatch	Player 1	20
2	Match	Mismatch	Player 2	20
3	Mismatch	Match	Player 1	20
4	Mismatch	Match	Player 2	20

Table 4.1. *Game-round structure played by pairs of participants during the Matching Pennies game*

4.3.1.5. Data coding and analysis

The experimenter live coded the data in both stages of the experiment. All analyses were conducted in R version 3.1.3. Generalised linear mixed models (GLMMs) were fitted using package lme4 (Bates et al., 2014), and significance testing was undertaken by likelihood-ratio tests (X^2) using the *anova* function. GLMMs were checked for overdispersion and to ensure that the magnitude of the standardized residuals were independent of the fitted values (Bolker et al., 2009).

4.3.2. Results

All pairs of participants allocated to the *cooperation encouraged* condition chose to work cooperatively on the buzz wire task, and all participants who were allocated to the *cooperation discouraged* condition chose to work individually. There was no difference in buzz wire performance between participants who chose to undertake the task cooperatively (mean= 6.8 wire runs), and those who completed the task individually (mean=5.9: $t_{(18.7)} = -1.02$, $p = .32$).

The percentage of matching gestures produced by the sighted player according to whether they were incentivized to match or mismatch their opponent's gestures, and by experimental condition, are shown in Table 4.2. An inspection of the percentages described in Table 4.2 indicates that the values are comparable with those presented in Belot et al. (2013), but provides no evidence that choosing to perform the buzz wire task cooperatively led to an increase in imitation: by contrast, gesture matching was lower in participants who had performed the task together, for games with both matching and mismatching incentives. Following Belot et al. (2013), one-tailed t-tests were performed on the percentage of matching gestures produced by

pairs of players, to determine if matching occurred at significantly greater levels than would be expected by chance (i.e., 50%). Across all responses combined, there was evidence that sighted players produced matching gestures significantly greater than expected by chance ($t_{(39)} = 2.04$, $p = .026$). However, there was no significant evidence of matching greater than chance when responses were considered separately for games with matching ($t_{(39)} = 1.61$, $p = .06$) or mismatching ($t_{(39)} = 1.13$, $p = .13$) incentives; a finding which appeared to result from a higher amount of variance existing within the measures when the data was split into separate response types.

A logistic regression, with standard errors adjusted for repeated measures on pairs of participants, was run to assess the effect of buzz wire undertaking (cooperatively/individually) on whether the players produced matching gestures. As expected from examination of the data in Table 1, there was no effect of condition ($X^2(1)=0.06$, odds ratio= 0.97, $p=.81$) on matching behaviour. Thus, in this experiment, there was no evidence that choosing to cooperate with another person led to an increase in automatic or intentional imitation.

	Performed buzzwire	
Game incentive	Cooperatively	Individually
Match	53.2%	53.9%
Mismatch	51.8%	52.5%

Table 4.2. *The percentage of game rounds in which the sighted player matched the gesture of the blindfolded player by game incentive (match/mismatch) and whether the player had chosen to perform the buzz wire task cooperatively or individually*

4.4. General Discussion

The three experiments presented in this chapter failed to find the expected positive relationship between cooperation and automatic imitation, despite participants reporting that they found their interaction partner cooperative (Experiment 2) and despite their choosing cooperation (Experiment 3). Experiment 1 produced results that hinted at an anticipated interaction effect – i.e. imitative behaviour was positively related to task performance when participants had interacted cooperatively, but was negatively related to performance when participants had performed alone – but this finding did not show any repeatability in Experiment 2, and is therefore seemingly not robust. This finding is reflective of growing concerns in psychology studies regarding the non-repeatability of results and effect sizes obtained with relatively small data samples (*Open Science Collaboration*., 2015), and should encourage caution when interpreting the results of similar studies that do not provide evidence of repeatability.

Experiment 3 attempted to address potential procedural weaknesses remaining from Experiment 1 and Experiment 2, by allowing participants the experience of 'choosing' to cooperate, and also employing a more accurate measure of imitation (gesture matching), yet this still failed to detect the expected increase in imitative behaviour within pairs who worked jointly relative to pairs who worked individually. On the contrary, pairs who worked jointly imitated each other at a lower rate than pairs who worked individually, though this effect did not reach statistical significance. This finding contrasts with that of Leighton et al. (2010), who did find an increase in automatic imitation of similar hand gestures (displayed via a computer monitor) in participants who had been primed with 'cooperative' versus 'individualist' words.

One possibility for the discrepancy between the findings in Experiment 3 and those of Leighton and colleagues is that joint performance on the buzz wire task

relative to performing the task individually does not instil the same differences in cooperative mind set as cooperative relative to individualist priming, even when participants have themselves chosen cooperation. For instance, the act of choosing to work individually in Experiment 3 might in itself have been experienced as a collaborative and beneficial act of decision-making between participant pairs, reducing the contrast between individual and cooperative task performance; though we should still expect decisions and experiences leading to cooperation to enhance cooperative rapport further than decisions to perform alone. It is also possible that joint decisions to perform the task individually led to a subsequent desire to reconnect and re-affiliate with the interaction partner (through increased imitation) during the Matching Pennies game (Lakin & Chartrand, 2003). Future research should repeat the procedure undertaken in Experiment 3 while replacing the buzz wire task with Leighton et al's cooperative priming procedure, to investigate these possibilities further.

An additional procedural difference that might also explain the difference in my own and Leighton et al's findings is that in the latter experiment participants were presented with a photographic image of the OPEN-HAND and CLOSED-HAND gestures via a computer display screen (a less naturalistic measure of automatic imitation that utilises a stimulus-response-compatibility/reaction-time paradigm; e.g., Brass et al., 2000; Bird et al., 2007), while participants in Experiment 3 were presented with the opportunity to imitate their own interaction partner's live hand gestures. However, this observation might lead us to expect a stronger relationship between cooperation and imitation in Experiment 3 relative to Leighton et al's measure, rather than the opposite pattern observed, as live measures of automatic

imitation (Belot et al., 2013; Cook et al., 2012) allow the imitator to respond directly to their cooperative or non-cooperative interaction partner's hand gestures.

As in the first two experiments, the discrepancy in findings between these two studies could also reflect the use of small data samples leading to random or systematic errors – 40 participants were used across two conditions in my study, while only 36 participants were used across three experimental conditions in the study by Leighton and colleagues. This possibility again highlights the need for replication in published research. Moreover, in my own experiment, given the very small percentage (%) differences in matching behaviour observed between experimental subcategories, and relative to chance matching, it is possible that even if the experimental manipulation were to produce an effect in the predicted direction, an unrealistically large sample would be required to detect what might accurately be a small effect size. Leighton et al's use of reaction time (i.e., the time taken by participants to produce a matching or mismatching gesture), rather than matching per say, to measure the effect of automatic imitation, may also represent a higher resolution and more amenable response variable in this regard.

It is also of course possible that the directional relationship leading from cooperative behaviour to imitation, which was the focus of study in this chapter, is not as strong as its reverse, or is mediated by other factors such as mood (van Baaren et al., 2006) or the degree of similarity between interaction partners (Van Swol & Drury-Grogan, *submitted*). This would explain the relative abundance of literature demonstrating the link between imitative behaviour and cooperativeness (e.g., Carpenter et al., 2013; Chartrand & Bargh, 1999; van Baaren et al., 2004), but the scarcity of empirical evidence showing the hypothesised relationship in the opposite direction. It is also plausible that anticipation of a cooperative interaction in the future

supports the relationship between cooperation and imitation (Lakin & Chartrand, 2003), while a previous cooperative interaction with no possibility of continued future cooperation (as in this experiment) does not. This possibility could also explain why cooperative priming produces the expected effect but previous cooperative behaviour did not. Future studies that vary the type of cooperation exercise and the context in which cooperation is required are necessary to investigate these possibilities further.

In sum, the experimental procedures detailed in this chapter did not find support for the relationship hypothesised between cooperative behaviour and imitation. They also highlight potential concerns relating to the repeatability and robustness of results obtained in investigations of automatic imitation, and provide support for the adoption of experimental procedures that demonstrate repeatability of results across more than one measure of imitation (e.g., face touching and foot tapping; Chartrand & Bargh, 1999).

Given the series of updated experimental manipulations adopted here, it is plausible that the relationship between cooperation and imitation is more nuanced than initially predicted, and likely mediated by other factors that were not represented in the experimental design. However, these results do not dispel the possibility that cooperation and imitation are interlocked in a mutually reinforcing system that bolsters cohesion and cultural transmission in social groups. Future studies are required to elucidate the conditions and contexts under which this relationship may operate.

Chapter 5

The Interaction Between Cooperation and Cultural Learning: An Experimental Investigation in Adults

Abstract

In the previous chapter, I investigated the relationship between cooperative behaviour and imitative behaviour, specifically attempting to assess the hypothesis that cooperation enhances the operation of mechanisms and behaviours leading to imitation. In this chapter, I attempted to examine the interaction between collective behaviour and instrumental cultural learning, using an experimental micro-society approach with groups of participants playing networked computer games. Despite an abundance of studies and theories purporting a link between large-scale cooperation and the evolution of complex culture (e.g., Chudek & Henrich, 2011; Hill et al., 2011; Pagel, 2012; Tomasello et al., 1993), there have been no attempts to manipulate this interaction experimentally. Specifically I tested whether individuals' access to cultural information in a learning game (as determined by other group members) was mediated by their reputation in a linked public goods game, and predicted that reputation-based access to information could serve to increase cooperative behaviour if free riders were targeted and discouraged by exclusion from cultural information. As predicted, I found evidence that participants preferentially shared information with co-operators and withheld information from free riders. However, the operation of reputation-based access to cultural information did not increase cooperative behaviour, relative to a control condition. It is likely that the cost of exclusion from cultural information was not large enough to offset the benefits of free riding in this experiment, but it is argued that this mechanism could serve to increase both cooperation and cultural exchange in the plausible scenario where being denied access to cultural information carries considerable costs.

5.1. Introduction

Laboratory-based economic games have been instrumental in generating empirical support for the paradigm that humans have evolved strong prosocial tendencies that bolster large-scale cooperation. The results of these experiments suggest that humans cooperate at much higher levels than is economically rational, even when there is no possibility of direct or indirect fitness benefits, immediately or in future (e.g., Fehr et al., 2002; Gintis et al., 2003; Rand et al., 2012; although for an alternative view see Burton-Chellew & West, 2013). The public goods game (PGG) is one such economic game that is frequently used to study cooperative behaviour in groups of humans faced with collective action decisions.

In the PGG, individuals choose how much to contribute to a public pool to produce benefits that are shared equally by all members of their group, including themselves. The maximum payoff for the group occurs when all individuals contribute the full amount. However, individual players can increase their own personal payoff by free riding – i.e. taking the benefits from the public pool while withholding their own contribution. Players are thus faced with a dilemma: the dilemma of contributing to the increased welfare and cooperativeness of the group, or of maximising their own immediate individual gain while risking retaliation from other group members in future. The PGG encapsulates the many public-goods scenarios that day-to-day engage us all: from reducing global carbon emissions, to adhering to hosepipe bans, to contributing fairly to community projects and household chores.

In a one-shot PGG, an individual wishing to maximise their personal income or gain should contribute zero. On the contrary, across a wide range of cultures, most players make initial contributions that are substantially greater than zero (~40-50% of

their resources; Gintis et al. 2003; Henrich et al., 2005; Burton-Chellew & West 2013). However, when play is repeated, the initial and unexpectedly high levels of investment by players in the PGG tend towards zero, as free-rider strategies emerge and players gradually switch to behaving in the expected self-interested manner (Fehr & Gächter 2000). This holds true whether the players (who interact anonymously via computer screens) are randomly assigned to groups in each round, or whether group composition, and therefore interaction partners, remain constant over multiple rounds (Fehr & Gächter 2000).

The mechanisms that sustain and bolster cooperative interaction in human groups, preventing the influx and establishment of free-rider strategies, have thus been the focus of much attention. The opportunity for game players to use costly punishment or reward has received particular focus. For example, in PGGs that permit individuals to pay a cost to punish non-contributors, cooperation does not deteriorate, but rather increases, over time (e.g., Fehr & Gächter, 2000; Gintis et al. 2003; Rand et al. 2009). Moreover, when interaction partners remain constant from round to round, costly reward (i.e., paying a cost to bestow a reward upon another group member) is found to be just as effective as punishment in maintaining high levels of cooperation, and can even lead to higher net payoffs (Rand et al. 2009).

In addition to costly punishment and reward, player behaviour in cooperative dilemmas is responsive to an increasing array of experimental manipulations. Symbolic sanctions (i.e., anticipating unrestricted verbal feedback from other players), that confer no cost on the sanctioner, can lead to increased cooperation, for instance (Ellingsen & Johannesson, 2008). Expectations of others' cooperativeness, which are higher when playing with members of one's 'cultural ingroup', might also influence game outcomes (Koopmans & Rebers, 2009). Likewise, conducting gameplay within

dynamic network settings, permitting players to change their interaction partners through making or breaking game connections, has also been shown to stabilise cooperation: co-operators frequently break their links with defectors whilst making new ties with other co-operators; discouraging freeriding and incentivising cooperation in the process (Rand et al., 2011). Moreover, the visibility of other players' wealth (game points) can lead to lower aggregate levels of cooperation within groups, and exacerbate initial levels of wealth inequality, relative to when wealth is hidden (Nishi et al., 2015).

However, despite the growing evidence of an intrinsic connection between human ultra-cooperativeness and cultural complexity (reviewed in Chapter 1: General Introduction), there has been very little experimental work that attempts to directly assess this interaction. Indirect evidence suggestive of a reciprocal and coevolutionary relationship between cultural exchange and the establishment of large, cooperative interaction networks, has been provided both by theoretical models (e.g., Henrich, 2004a; Powell et al., 2009; Pradhan et al., 2012), and empirical studies (e.g., Discamps & Henshilwood, 2015; Henrich et al., 2005; Hill et al., 2011, 2014). One experimental study, comparing the learning capabilities of children, chimpanzees and capuchins, supports the suggestion that human cultural feats, such as cumulative technological advancements, are directly scaffolded by our extreme cooperativeness (Dean et al., 2012). Yet, to my knowledge, there have been no attempts to experimentally examine whether our extreme dependency on cultural information in turn promotes and scaffolds cooperation. Because cooperative interactions and cultural exchange have likely coevolved (Chudek & Henrich, 2011; Hill et al., 2011; Tomasello, 1999), it is plausible that our requirement for up-to-date and reliable

cultural information acts to encourage and stabilize cooperative behaviour, in addition to depending on it.

In this chapter, I attempted to study the relationship between cultural transmission and cooperative behaviour experimentally, using an experimental micro-society approach (Mesoudi & Whiten, 2008). Specifically, I investigated whether individuals' reputations (i.e, contributions) in a public goods game influenced their access (as determined by fellow group members) to beneficial 'cultural' information, in a separate but linked learning game. In turn, this also permitted me to examine whether reputation-based access to cultural information could encourage increased cooperative reputation building, thereby serving to promote and increase group-level cooperation. In this experiment, following each round of an iterated PGG, all group members learnt of each other's contribution to the PGG then each played one round of a separate computer-based learning game. During each play of the learning game, individuals were permitted to choose which other group members they were willing to share game-related ('cultural') information with. In a second experimental condition, that prevented and therefore controlled for reputation-based effects across the two games, players were assigned separate and *unlinked* identities in the PGG and learning game.

For the learning game, I created an adapted version of a custom-made cumulative learning game, 'the monster league game', previously implemented by Wisdom et al. (2013) to study social learning in networked groups. This game, which is outlined in detail below (see Methods), requires players to select a team of monster icons from a league of available icons, and players aim to maximise the score of their selected team in each round. The hidden score distribution in the monster league game prevents individual participants from monopolising increments in performance, and

thus encourages continued motivation to engage in information exchange among players. Moreover, a previous study (Wisdom et al. 2013) showed that groups with access to each other's game play were able to explore the problem space more effectively than groups with reduced social information, and that the advantage of sharing information was also positively correlated with the number of group members contributing information.

When participants' identities were linked across the PGG and learning game, they were expected to preferentially direct information sharing towards co-operators, and withhold information from free riders. However, when participant game identities were unlinked, this reputation-based specificity in information sharing was prevented. Thus, I predicted that higher levels of cooperation (i.e., PGG contributions) would be maintained when game identities were linked, as reputation-based targeted exclusion from social information would act to discourage individuals from free riding in this condition, but would be prevented from operating when game identities were unlinked.

In turn, I predicted that, when game identities were linked, group-level PGG contributions would be positively related to information transfers, and individual-level PGG contributions would be positively related to the amount of information transfers individuals received. No prediction was made regarding the overall relationship between PGG contributions and information transfers in the unlinked identity condition, nor regarding overall differences in information transfers between conditions, as there was no clear hypothesis regarding the absolute level of information transfer that would occur when participants' between-game identities were unlinked. However, relatedly, I also anticipated that, within experimental conditions, participants might be relatively less willing to share high-quality

information (i.e., high-scoring monster teams) when they were unsure of other individuals' reputations (i.e., when game identities were unlinked), than when the reputations of their information recipients were visible (i.e., game identities were linked).

In line with previous findings (Wisdom et al. 2013), I predicted that learning game performance would be positively associated with the amount of social information transfer between group members, at the group and individual level. I did not make a firm prediction regarding differences in learning game performance between conditions, due to the unpredictability of between-condition differences in information transfer. However, I did not expect that learning game performance in the unlinked identity condition would exceed that in the linked identity condition (although it could be equal or less), due to the expectation that positive feedback between cooperation and sharing when identities were linked would drive high levels of performance. I also predicted that PGG contributions would be positively associated with learning game performance when game identities were linked, due to the positive relationship expected between PGG contributions and information transfer in this condition, but that this relationship would not be observed when game identities were unlinked, due to the inability of participants to link PGG reputations to information-transfer decisions in this condition. Finally, I anticipated that participants' game-round payoff (i.e., total points earned across the PGG and learning game in each round) would be higher when game identities were linked than when they were unlinked, due to the higher and more stable levels of cooperation expected in this condition.

5.2. Methods

5.2.1. Participants and materials

Participants were 80 students and staff at the University of St Andrews, recruited via the University's online sign-up system for human experiments (SONA). Participation was voluntary, and participants were remunerated with a £5 participation fee, plus the opportunity to earn an additional bonus of up to £15 dependent on their performance. Participants were aged between 18 and 52 years old (mean=21.5, sd=4.8), and 32 were male. There was no difference in the distribution of age (Mann-Whitney: W=762.5, p=0.72) or sex ($X^2(1)$=0.47, p=0.49) between experimental conditions. Ethical approval for this study was given by UTREC of the University of St Andrews.

All participants took part in groups of four, and each experimental session was conducted inside a university computer lab with one or two groups of four participants (i.e., 4 or 8 participants per session). Each participant had access to a computer, and was visually separated from the other participants by large screens. The experiment was implemented via custom-made software written in Java, and run over a local computer network. Participants used a mouse and keyboard to interact with the experiment, and participants' computers were coordinated and updated by a server computer, which also recorded the data, throughout the experiment.

5.2.2. Computer games

Public goods game (PGG)

Following Rand et al. (2009), I employed a standard PGG design (designated *the project game*), in which groups of four participants interacted repeatedly. In every game round, each participant received 20 monetary points and was asked to decide

how many points (between 0 and 20) they were willing to contribute to the public good (i.e., the group project), on the proviso that they would keep any points they did not contribute for themselves. Participants were given 20 seconds to make their contribution, and were informed that failure to decide in this timeframe would result in a random number of points between 0 and 20 being contributed. Contributions to the public good were then doubled before being split evenly back among the four group members.

Thus, the maximum payoff for the group occurred when all participants contributed 20 points, and each participant received back 40 points (i.e., for each point contributed, all participants received back 0.5 points), but individual participants could increase their own game-round earnings relative to other group members by contributing less while still benefiting from the group pool (the maximum individual payoff was 50 points for an individual who contributed zero when all other group members contributed 20 points).

Learning game

The learning game (designated the *monster league game*) was adapted from, and closely resembled, a custom-made learning game designed by Wisdom et al. (2013). The aim of the game for each participant was to maximise the number of points earned by a subset of six monster icons ("monster team") chosen from a larger subset of 48 monster icons ("the monster league"). The game screen included the league of 48 monster icons, an area for the participant's current team selection, and another area displaying the participant's best-scoring team alongside their team from the previous round (see Figure 5.1 for a screen shot of the learning game). The scores of the participant's best-scoring and previous round teams were also visible, and a timer

located in the top right corner of the screen reminded participants of how many seconds remained in the round. In each round of the game, participants were given 20 seconds to choose six icons for their current monster team, after which random monster icons were selected from the available icons in the league to fill any remaining spaces. Participants could click on a button to "clear" their current team entirely, or alternatively could clear individual monster icons they wished to change by clicking on the icon. Monster icons could be selected from any part of the screen by clicking on the icon, except those icons that already appeared in the participant's current team which were highlighted with a green border. In addition, the current team could be replaced entirely by a previous team through clicking on the "copy" button that appeared beneath it.

Figure 5.1. *Screen shot of the learning game during the practice rounds. During interactive gameplay participants could also view the monster teams of other participants who had agreed to share with them (see Figure 5.3)*

In each new experimental session, the league and monster icon positions were generated randomly from a larger set of 149 monster icons, but then remained fixed for the duration of the session. Following Wisdom et al. (2013), each monster icon was associated with a specified positive points score, and some 'interaction' pairs of monster icons were associated with additional points bonuses or penalties when they appeared in the same team together. Thus, the score for each monster team represented the sum of points associated with each monster icon plus or minus the value of any interaction pairs that were present. There was no overlap between interaction pairs, and their distribution, which concentrated bonuses among low-scoring pairs and penalties among high-scoring pairs, was previously found to be challenging for experimental participants (Wisdom et al. 2013; see Figure 5.2a for an outline of the points distribution).

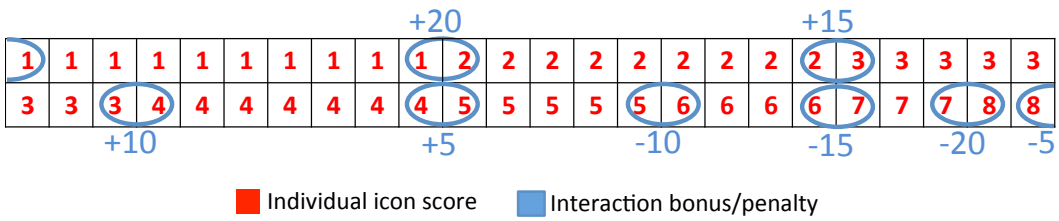


Figure 5.2a. *The score distribution applied to individual monster icons and the interaction bonuses and penalties applied to pairs of icons, for a league size of 48 monster icons (adapted from Wisdom et al., 2013).*

Participants were provided with information about the maximum possible team score achievable (i.e., 60 points), to motivate continued striving for high scores throughout repeated rounds of the game, and were informed that some pairs of

monster icons conferred bonuses or penalties, but were not given any information about the score distribution (see Figure 5.2b for the probability distribution of scores among all possible teams).

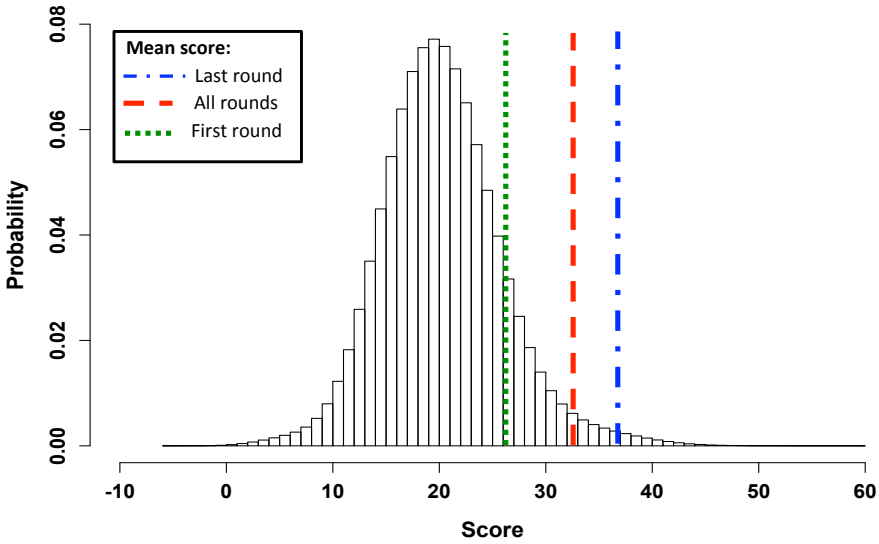


Figure 5.2b. *The probability distribution of monster team scores [-6,60] for a league of size 48 and the score distribution displayed in Figure 2a. The mean score in the first round (green line), last round (blue line), and overall (red line) for all participants is also shown.*

Participants played the learning game in the same group of four participants with whom they played the public goods game. In each round, participants could choose which of the other three participants they wished to share their previous round monster team and score with, and were permitted to copy all, some or none of the solutions that the other participants had chosen to share with them, which were displayed in a separate area of the game screen (see below, section 5.*2.3. Design*, for

further details of participant interactions in the learning game across experimental conditions).

5.2.3. Design

In a between-groups design, I randomly allocated groups of four participants to one of two experimental conditions. Participants were informed they would interact with the same three people throughout the experiment. In the first condition (the *linked identity* condition, N=40), each participant was assigned a single identity at the beginning of the experiment (e.g., 'Player 1'), which they kept in both the PGG and learning game for the duration of the experiment. In contrast, participants allocated to the second condition (the *unlinked identity* condition, N=40) were allocated separate identities for each game (e.g., 'Player 1' in the PGG and 'Player Blue' in the learning game), which also remained fixed for the duration of the experiment, but were not linkable between games. Participants interacted anonymously via their computer stations (it was not possible to link the player identities in the games to the other people in the room).

At the beginning of each session, participants then undertook an extensive computer-based tutorial in the rules and mechanics of the learning game (known to players as the *monster league game*) followed by the PGG (known to players as the *project game*), before beginning iterative rounds of interactive gameplay with group members (see section 5.2.2. for a description of the games). During the game tutorials, participants were also instructed that individual bonus payments of up to £15 would be calculated at the end of the experimental session, and would consist of their cumulative points earned in both games following a conversion of: 145 points = £1.

The learning game tutorial included five practice rounds of the game, which participants played individually. Participants were instructed that the scores they

obtained in the practice rounds would not contribute to their final payment, but that they would carryover their score and game information (i.e., monster team) from the final practice round into the first round of the interactive game. Following Rand et al. (2009), the public goods game tutorial included detailed instructions about the game rules, including a series of worked hypothetical examples, but did not involve a practice period. Participants were permitted to complete the game tutorials in their own time, with those who completed early held in a game 'waiting room' until all other group members had completed.

Participants in the *linked identity* condition were informed that in each game round, one play of the public goods game would always be followed by one play of the learning game. Participants were also aware that they would learn of each other's public goods contributions before each play of the learning game, and would have to decide with whom they wished to share information from their last attempt at the learning game (i.e., their monster team and score from the previous round) in the next play of the game (see Figure 5.3a). Participants in the *unlinked identity* condition received the same instructions, with one important exception. After participants in the *unlinked identity* condition had learned of each other's contributions to the PGG, they were reassigned their separate learning game identities, and had their screen positions randomly shuffled, before deciding whom they wished to share their last round's monster team with (i.e., individual identities and reputations gained in the public goods game could not be linked to identities and directed-sharing decisions in the learning game, and vice versa; see Figure 5.3b).

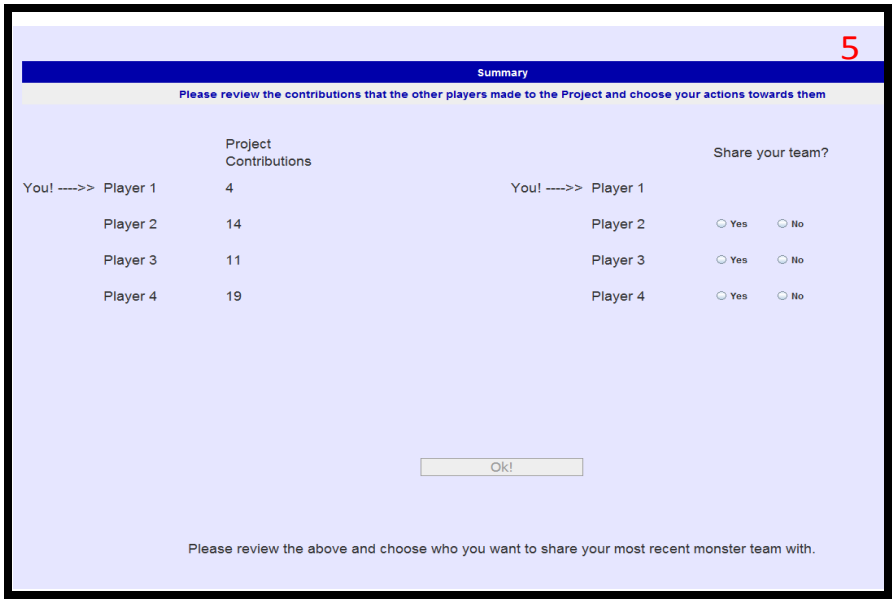


Figure 5.3a. *Screenshot example of decision-making screen as seen by participants in the linked identity condition in each round prior to playing the learning game*

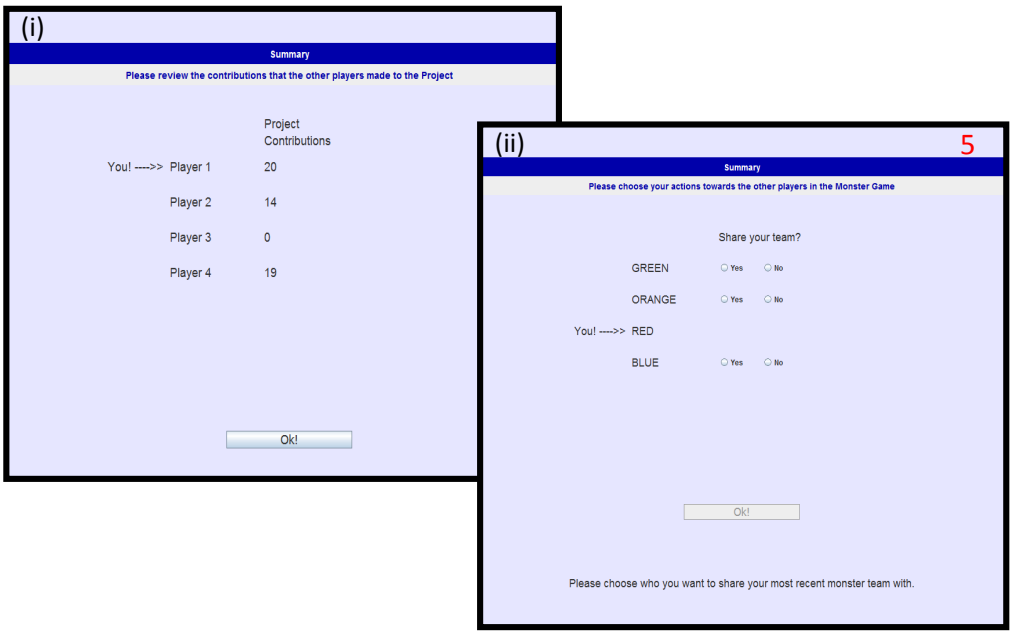


Figure 5.3b. *Screenshot examples of public goods game results screen (i), which was immediately followed by identity reassignment and the learning game decision-making screen (ii), as seen by participants in the unlinked identity condition in each round prior to playing the learning game*

Participants were given 20 seconds to decide which of the other players they would share information with, before playing one round of the learning game. While playing the learning game, participants had visual access to the scores and monster teams submitted in the previous round by group members who had agreed to share with them, but the scores and monster teams of group members who had not agreed to share remained hidden (see Figure 5.4 for a screenshot example from the *linked identity* condition). Players could copy some or all of their own monster team from the information provided by other group members, but were also free to ignore this information.

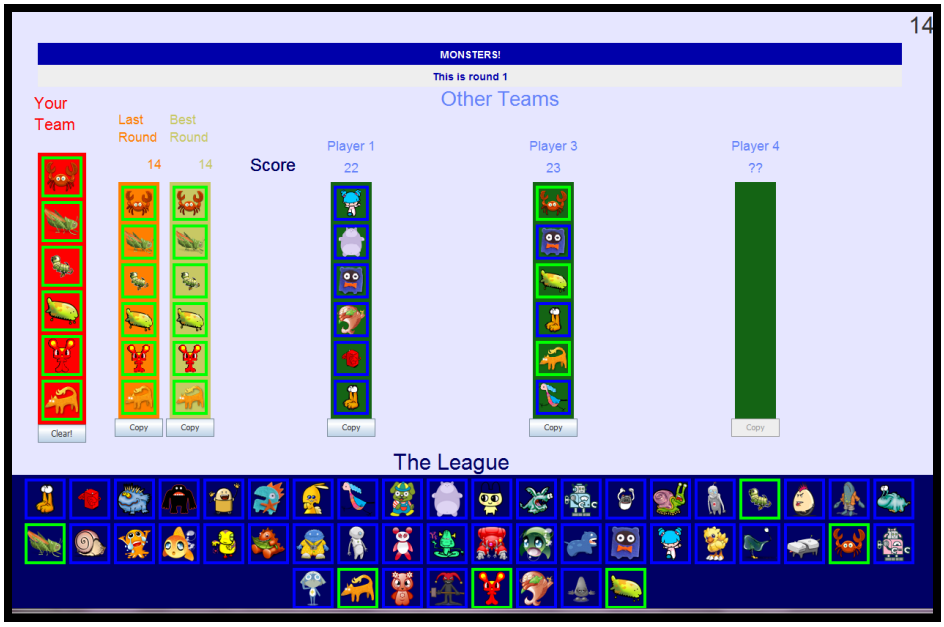


Figure 5.4. *Screenshot example of the interactive learning game in the linked identity condition (Screen as seen by Player 2: Player 1 and 3 agreed to share their previous team with Player 2, but Player 4 withheld information in this round.)*

At the end of every round, participants saw a summary of the round, reminding them of the contributions made by each player to the public good, and of which players had shared information with them. The summary also detailed the participant's current round points and cumulative points separately for the public goods game and learning game. Participants then began the next round of game play until 20 rounds had elapsed, after which the session ended (20 rounds was chosen as Wisdom et al. (2013) previously reported a performance plateau in in the learning game after 24 rounds). Participants were informed that they would play an unspecified number of rounds, and that their behaviour during the game would not affect the total number of rounds played, so were unaware of when the game would end (see Figure 5.5 for a schematic of the game structure).

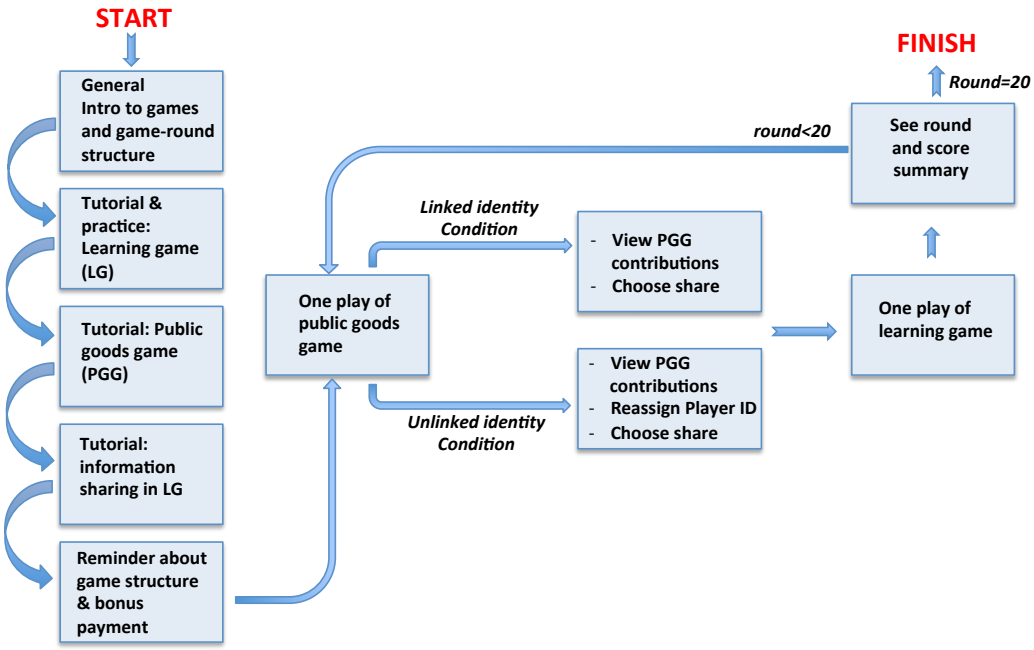


Figure 5.5. *Schematic of the game stages and progression of participants through iterative rounds of the public goods game (PGG) and learning game (LG), in the linked identity and unlinked identity conditions*

5.2.4. Procedure

Participants were recruited for a study on 'Interactive decision-making,' and those who were known to each other were prevented from participating in the same experimental session. On arrival, participants were given a verbal and written briefing by the experimenter, and provided with an opportunity to ask questions. The experimenter then started the experiment on participants' computers remotely from a server computer located in an adjoining room. The experimenter remained in the adjoining room with the connecting door open for the duration of the experiment, to ensure that no participants communicated. There were no instances of between-participant communication observed. The experiment lasted approximately 45 minutes, and after it had ended the experimenter returned to debrief participants and answer any remaining questions. Participants were then free to collect their payment and leave.

5.3. Analyses

All analyses were conducted in R version 3.1.3. Linear mixed models (LMMs) and generalised linear mixed models (GLMMs) were fitted using package lme4 (Bates et al., 2014).

LMMs were used when the dependent variable measured the public goods contribution or the number of points scored in a given round (group means were modelled in analyses that investigated outcomes at the level of groups). Model standardised residuals were examined for deviations from homoscedasticity and normality when LMMs were used. However, GLMMs with a binomial error structure

were employed when the dependent variable represented the proportion of agreed information transfers that occurred within the learning game. At the group level, a maximum of 12 information transfers could take place between groups of four participants in each game round, and the proportion of agreed information transfers represented the observed number of transfers divided by 12. In analyses at the individual level, the maximum number of information transfers a participant could both provide and receive was 3, therefore the proportion of information shares an individual received or provided had a denominator of 3. All GLMMs were checked for over-dispersion and to ensure that the magnitude of the standardized residuals were independent of the fitted values (Bolker et al., 2009).

For analyses that investigated behavioural decisions and outcomes at the individual level, I fitted "individual ID" and "group ID" as random intercepts (factors) to account for the non-independence of measures on individuals and groups across repeated rounds. When analyses were undertaken at the group level, I fitted "group ID" as a random intercept. In all models, a fixed effect was included for round number. Additionally, each random intercept was also fitted with an independent random slope for the effect of round number, to account for possible variation in the slopes of individuals and groups across rounds.

Unless otherwise stated, analyses of the data collected within each of the games (i.e., within the public goods game *or* the learning game) were conducted on the aggregate dataset. Analyses involving data collected across the two games were conducted separately for each experimental condition, due to differences in between-game dynamics between conditions. Significance testing was undertaken by likelihood-ratio tests (LRTs) using the *anova* function. LRTs compare the change in

residual deviance (which follows a X^2 distribution) that occurs between nested models following the removal/addition of a variable.

5.4. Results

The results are presented in six sections examining: (1) average contributions to the public good; (2) information transfers in the learning game; (3) the relationship between public goods contributions and access to social information, at the (i) group and (ii) individual level (*linked identity* condition only); (4) performance on the learning task, and the influence of (i) information access and (ii) cooperative behaviour on learning task performance (at both the group and individual level); (5) the effect of participants' learning task performance on their willingness to share information with other group members; and, (6) the relationship between participants' contributions to the public good and their total game-round payoff.

5.4.1. Contributions to the public goods game

Contrary to expectations, mean overall game contributions were similar across conditions: 11.8 (±SE 0.26) in the *linked identity* condition and 11.6 (±SE 0.25) in the *unlinked identity* condition. Thus, making visible the PGG reputations of individuals during decisions about information transfer did not lead to increases in PGG contributions (Group level analysis: LMM: $X^2(1)=0$, $p=.96$, N=400; see Figure 5.6). In both the *linked*- and *unlinked identity* conditions, contributions decreased slightly over trials (Mean 1st round contribution: 13.0; last round contribution: 10.3; LMM: $X^2(1)= 9.17$, $p=.002$)

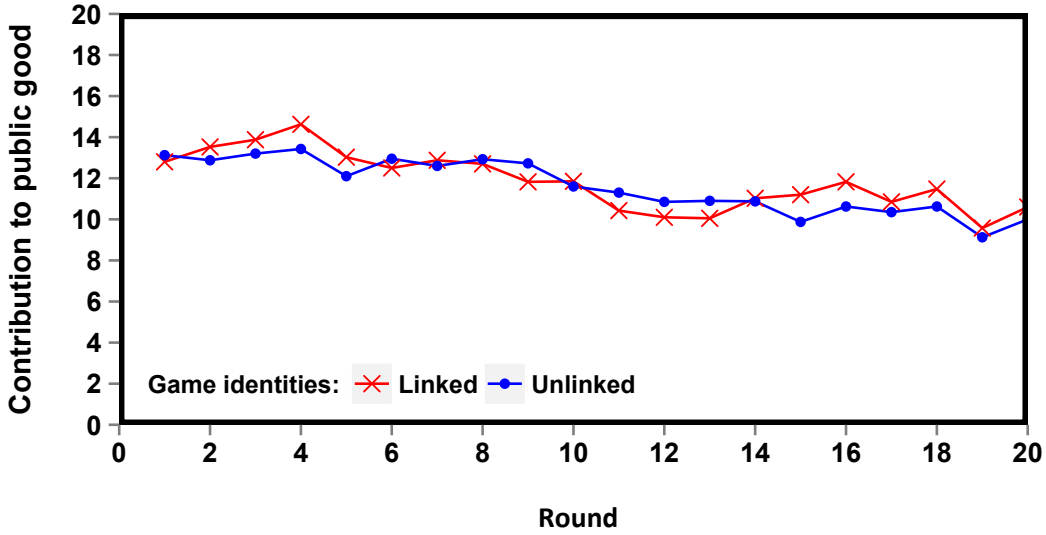


Figure 5.6. *Average contribution to the public good by round number and condition*

5.4.2. Information transfers

An analysis to examine the effect of experimental condition on information transfers was conducted at the group level. There had been no prior predictions made about between-condition differences in information sharing (see 5.1. Introduction). Participants in the *linked identity* condition withheld information more frequently than participants with *unlinked* game identities (GLMM: $X^2(1)= 10.04$, odds ratio=3.98, p=.002, N= 400; Figure 5.7). Thus, making participants' PGG reputations visible within the learning game reduced overall levels of information transfer, despite extremely similar levels of PG contributions between the two conditions. There was no significant effect of round number on information sharing across conditions (GLMM: $X^2(1)= 0.90$, odds ratio=0.98, p=.34 , N= 400).

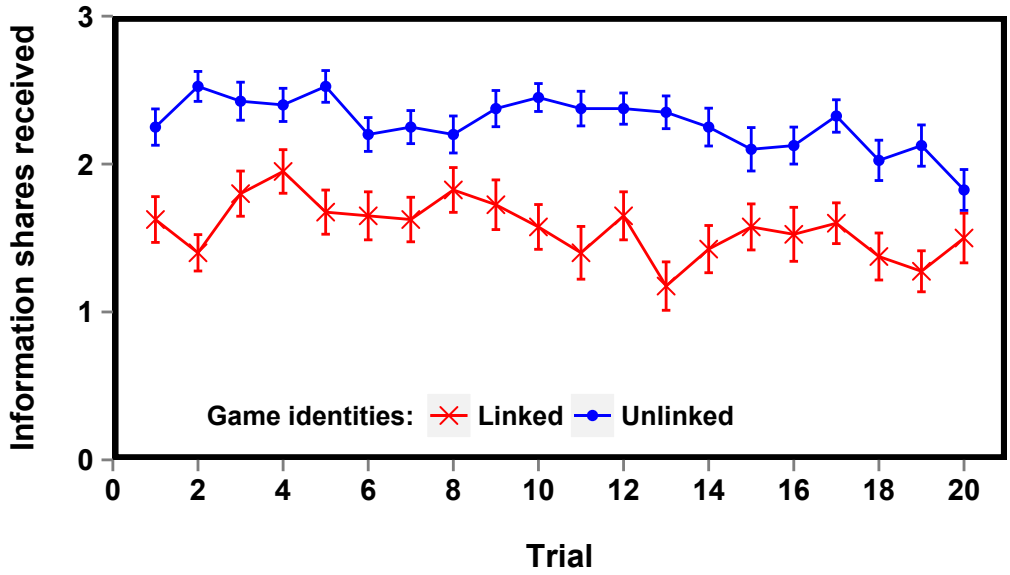


Figure 5.7. *The average number of information shares that participants received from group members by round number and condition ±SE (the maximum number of shares an individual player could receive in each round was 3)*

5.4.3. Public goods contributions and information transfers

The relationship between cooperative behaviour in the PGG and information transfers was examined at the *group* level for each condition separately. I then examined whether more cooperative individuals received greater access to social information within groups, using data collected in the *linked identity* condition only (participants' with *unlinked* game identities were excluded from this analysis as individual PGG reputations were not discernible during learning game decision-making in this condition).

Group-level contributions and information transfers

As predicted, there was a significant positive relationship between average group contributions and information transfers in the *linked identity* condition (GLMM:

$X^2(1)=107.77$, odds ratio= 1.2, $p<.001$, N=200; see Fig 5.8a). A weaker positive relationship was also observed in the *unlinked identity* condition (GLMM: $X^2(1)=14.99$, odds ratio= 1.1, $p<.001$, N=200; see Fig 5.8b), although no prior prediction had been made regarding the relationship between contributions and information transfers in this condition. There was no effect of round number on information transfers in either condition (p>.05).

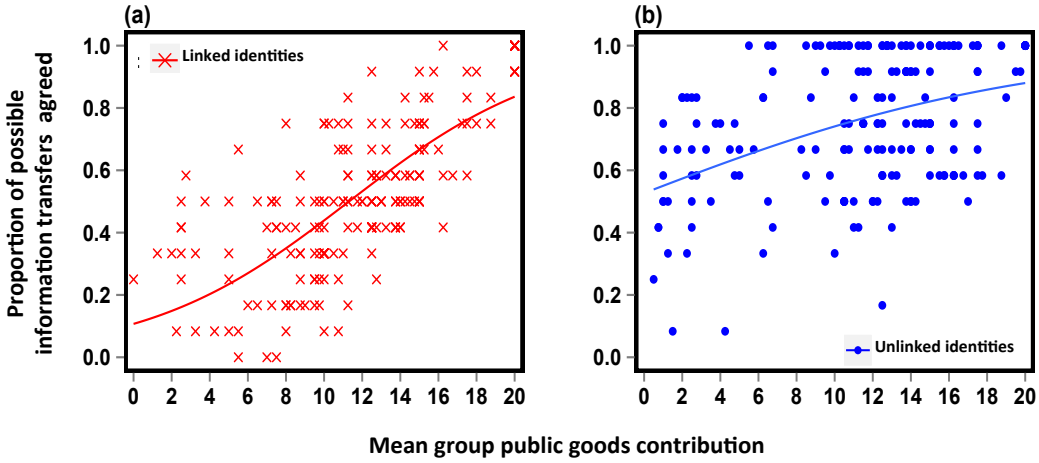


Figure 5.8. *Scatter plots showing the average contribution to the public good and the proportion of all possible information transfers agreed in (a) the linked identity, and (b) the unlinked identity condition. (Dots represent group-level raw data values with binomial lines of best fit.)*

Individual contributions (reputations) and access to shared information

The *linked identity* condition was examined to assess the effect of individuals' contributions to the public good on the proportion of group members who were willing to share information with them. As participants' sharing decisions might show greater sensitivity to the relative rather than absolute contributions of other group

members, analyses were conducted using both absolute and standardized contribution measures (standardized contributions were computed as the number of standard deviations of each participant's contribution from the group mean in a given round). Standardized individual scores could not be computed when all members of a group contributed equally to the public good, therefore data was excluded from analyses at the group level in rounds where there was zero variation in contributions (80 from a total 800 observations were removed, leading to analyses conducted on 720 individual observations). The amount of information a participant received might also have been influenced by the amount of information they had previously shared with other group members (i.e., a reciprocal process of information exchange might exist within the learning game itself), therefore an additional variable detailing the number of group members the participant had shared with in the previous round was included in the model.

Although both measures were statistically significant, the absolute value of participants' public goods contributions explained a larger amount of variation in the number of information shares they received than their standardized contributions (ΔAIC= -77), suggesting that absolute rather than relative contributions were more influential in decision-making. In line with my predictions, participants who made higher contributions received a greater number of information shares from fellow group members (*absolute contributions*: GLMM: $X^2(1)=207.69$, odds ratio= 1.13, $p<.001$, N=720; see Figure 5.9). The number of information shares participants received was not significantly affected by the amount of information they themselves had shared in the previous round, nor by round number, after accounting for their contribution to the public good ($p>.05$).

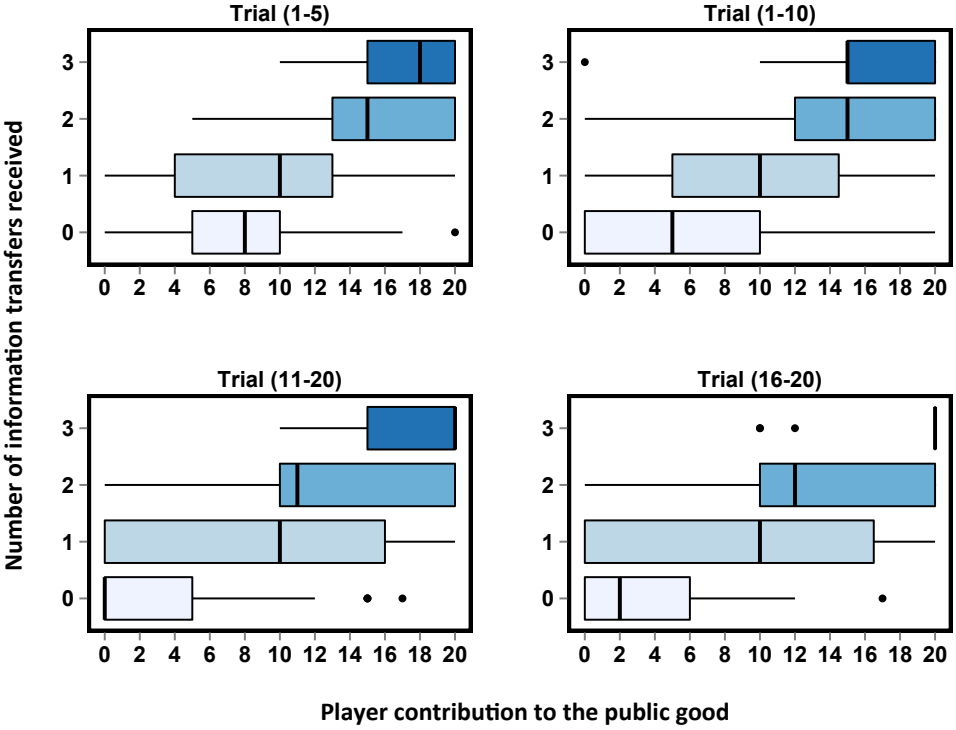


Figure 5.9. *Boxplot of the relationship between participant contributions in the linked identity condition and the number of information transfers they received from group members (out of a maximum of 3), at different stages (rounds) in the game*

5.4.4. Performance on the learning task

Performance on the learning task was first examined for differences between conditions. Subsequently, I examined the effects of information transfers and cooperative behaviour on learning task performance, at the level of both groups and individuals.

Learning task performance between conditions

No *a priori* predictions were made regarding differences in performance on the learning task between conditions. However, in line with previous findings (Wisdom et al., 2013), I did expect that higher rates of information transfer would be positively

associated with performance. Despite participants in the *unlinked identity* condition engaging in significantly higher levels of information transfer than participants in the *linked identity* condition (see section 5.*4.2.*), there were no between-condition differences in learning game performance (Group level analysis: LMM: $X^2(1)=0.07$, $p=0.79$, N=400; mean score *linked identity*: 32.6; mean score *unlinked identity*: 32.8; see Figure 5.10), implying that higher rates of information transfer were not advantageous at the aggregate level in this experiment. There was a significant, positive effect of round number across both conditions, as scores showed cumulative improvements over time (Group level analysis: LMM: $X^2(1)=28.72$, $\beta=0.44$, $p<.001$, N=400; mean score first round: 26.2; mean score last round: 36.6).

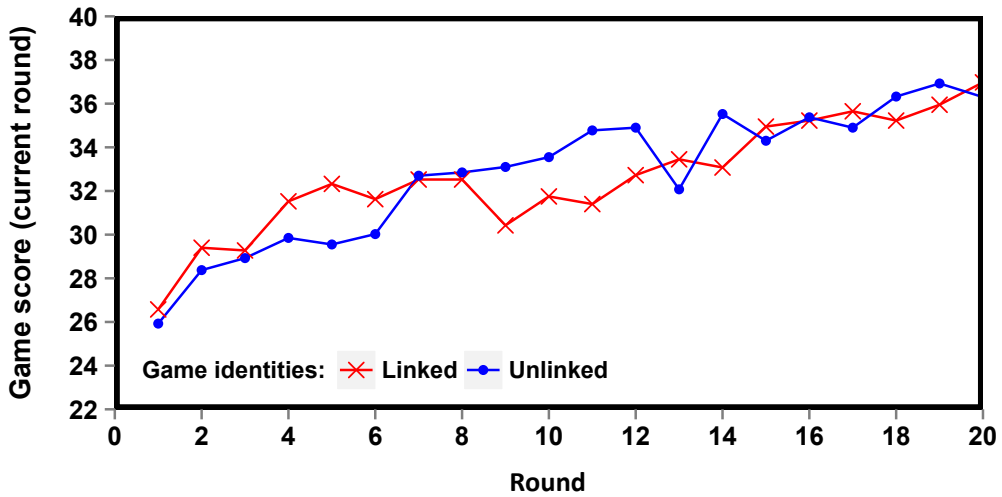


Figure 5.10. *Average game-round score by round number and condition. The mean score in the linked identity condition was 32.6 [min=1; max=48], and was 32.8 [min=4; max=47] in the unlinked identity condition*

Information transfers and learning task performance

To examine the effect of information transfer on learning game performance in more detail, analyses were conducted separately for each condition to allow for the possibility of different within-condition dynamics in the variables of interest.

At the group level, while controlling for round number, a positive relationship was found between information transfers and learning task performance in the *linked identity* condition (LMM: $X^2(1)=8.55$, $\beta=0.31$, $p=.003$, N=200; Figure 5.11a), but not in the *unlinked identity* condition (LMM: $X^2(1)=0.39$, $\beta=-0.09$, $p=.53$, N=200; Figure 5.11b). Thus, the expected positive association between information transfers and learning task performance was confirmed only in the *linked identity* condition. A similar pattern of results was observed across conditions when analyses were conducted at the level of individuals (LMM: *linked identity:* $X^2(1)=7.45$, $\beta=0.6$, $p=.006$, N=800; *unlinked identity:* $X^2(1)=0.27$, $\beta=0.17$, $p=.61$, N=800).

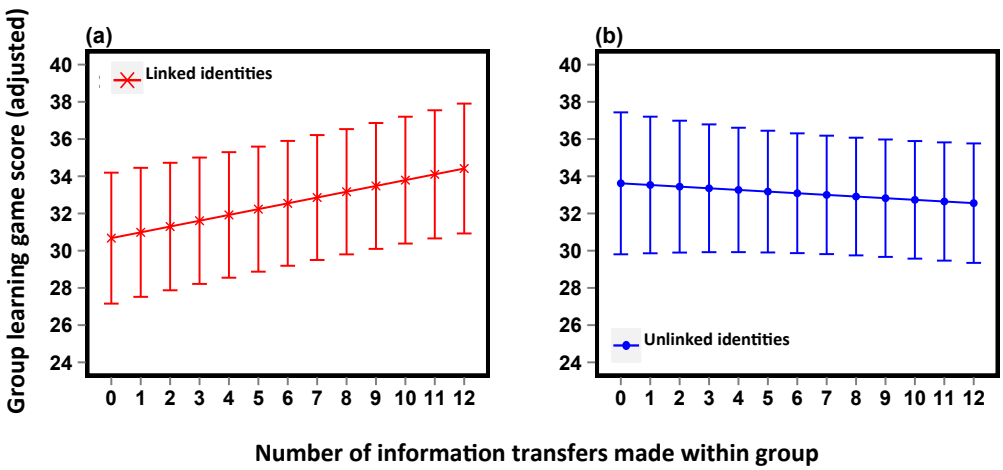


Figure 5.11. *Predicted relationship between a group's mean learning game performance and the number of within-group information transfers in a given round (±95% CIs), for the (a) linked identity condition, and (b) unlinked identity condition (adjusted for the effect of game round)*

Cooperative behaviour and learning task performance

Section 5.4.3 showed that, at the group level, the positive relationship between cooperative behaviour and social information transfer was stronger in the *linked identity* condition than in the *unlinked identity* condition. An analysis conducted at the group level, controlling for round number, also confirmed a significant and positive relationship between cooperation and learning task performance in the *linked identity* condition (LMM: $X^2(1)=7.46$, $\beta=0.22$, $p=.006$, N=200; Figure 5.12a), but not among groups of participants whose identities were *unlinked* (LMM: $X^2(1)=0.22$, $\beta=0.03$, $p=.64$, N=200; Figure 5.12b). Thus co-operators fared better in the learning game when their reputations were revealed, but not when their reputations were hidden, supporting the earlier predictions. Data was also analysed at the individual level for the *linked identity* condition only, and showed a similar, positive relationship between cooperative behaviour and learning task performance (LMM: $X^2(1)=0.22$, $\beta=0.11$, $p=.008$, N=800).

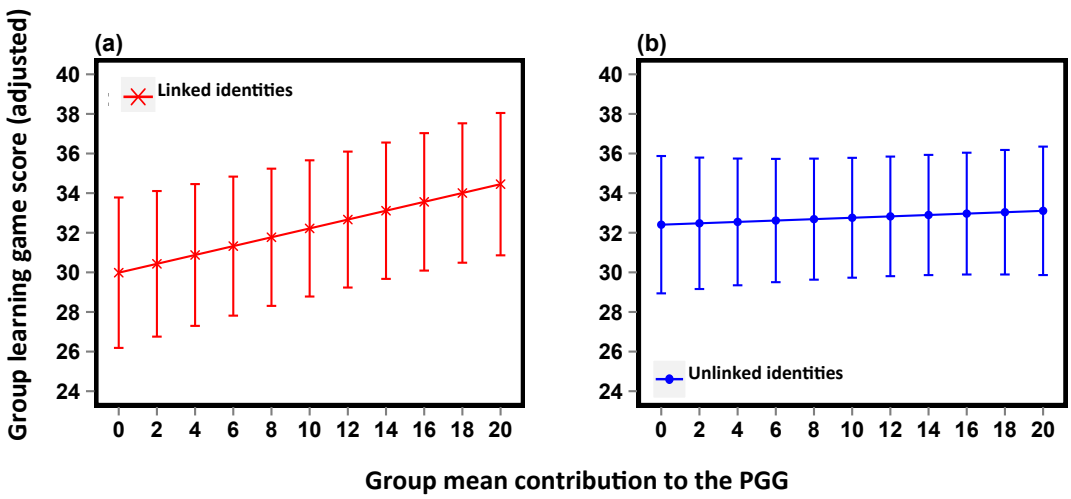


Figure 5.12. *Predicted relationship between a group's mean public goods game (PGG) contribution and its mean learning game performance in a given round (±95% CIs), for the (a) linked identity condition, and (b) unlinked identity condition (adjusted for the effect of game round)*

5.4.5. Performance on the learning task and decisions to share information

To investigate the prediction that participants would be relatively more willing to share high-quality information (high-scoring monster teams) when the cooperative reputations of the information recipients were visible, an analysis was conducted at the individual level, which included an interaction term between condition and the score of the monster team available for sharing. Variables detailing the mean current round contribution made by other group members, and the number of shares received from other group members in the previous round, including their respective interactions with experimental condition, were also included in the full model, to control for possible confounding effects from these variables (see *Section 5.4.3. & Section 5.4.4.*).

As predicted, there was a significant interaction between experimental condition and the quality of information shared, suggesting that participants in the *linked identity* condition were more likely to share high-quality information than participants in the *unlinked identity* condition, after accounting for between-condition differences in sharing behaviour (GLMM: $X^2(1)=8.45$, $p=.004$, N=1600; see Figure 5.13). The average public goods contribution made by other group members also exerted a positive main effect on participants' decisions to share information (GLMM: $X^2(1)=143.01$, odds ratio=1.16, $p<.001$, N=1600), but did not show an interaction effect with condition ($p>.05$). The number of information shares received by group members in the previous round did not significantly influence participants' decisions to share information, nor show an interaction effect with condition ($p>.05$).

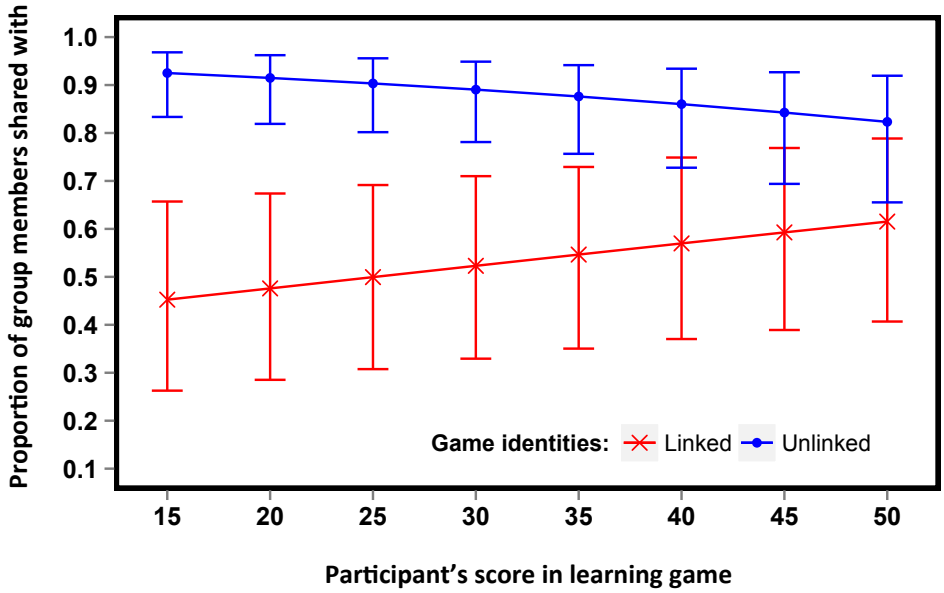


Figure 5.13. *The predicted proportion of group members shared with (±95% CI) by condition and the participant's previous-round score in the learning game (adjusted for the effect of game round and the average public goods contribution made by fellow group members)*

5.4.6. Total game-round payoff relative to public goods contribution

The preceding results sections provided evidence that PGG contributions were positively linked with individuals' access to social information and individual performance in the learning game when participants maintained linked identities and reputations between games. However, unexpectedly, the benefits conferred to cooperative individuals in the *linked identity* condition did not appear to garner increased or sustained levels of public goods contributions relative to a condition where participants' identities and reputations were unlinked between games. This is further reflected by the failure to find support for the prediction that individuals' game-round payoff would be higher in the *linked identity* condition than the *unlinked*

identity condition (LMM: $X^2(1)=0.0$, $p=0.98$, N=1600; mean game-round score *linked identity*: 64.3; *unlinked identity*: 64.2).

One possibility for this finding is that the cost bestowed on free riders in the public goods game was not sufficient to offset the benefits gained, and that individuals were perceptive to this (i.e., individuals were still incentivised to free ride). To examine this possibility, I analysed the relationship between participants' standardised public goods contributions and their total game-round points (i.e. payoff), for each of the conditions separately. Standardised public goods game contributions were found to negatively correlate with game-round payoff in both the *linked identity* (LMM: $X^2(1)=697.49$, $\beta=-4.85$, $p<.001$, N=720; Figure 5.13a) and the *unlinked identity* conditions (LMM: $X^2(1)=523.03$, $\beta=-4.32$, $p<.001$, N=748; Figure 5.13b). Thus, across both conditions, individuals still benefitted overall from making lower public goods contributions than other group members.

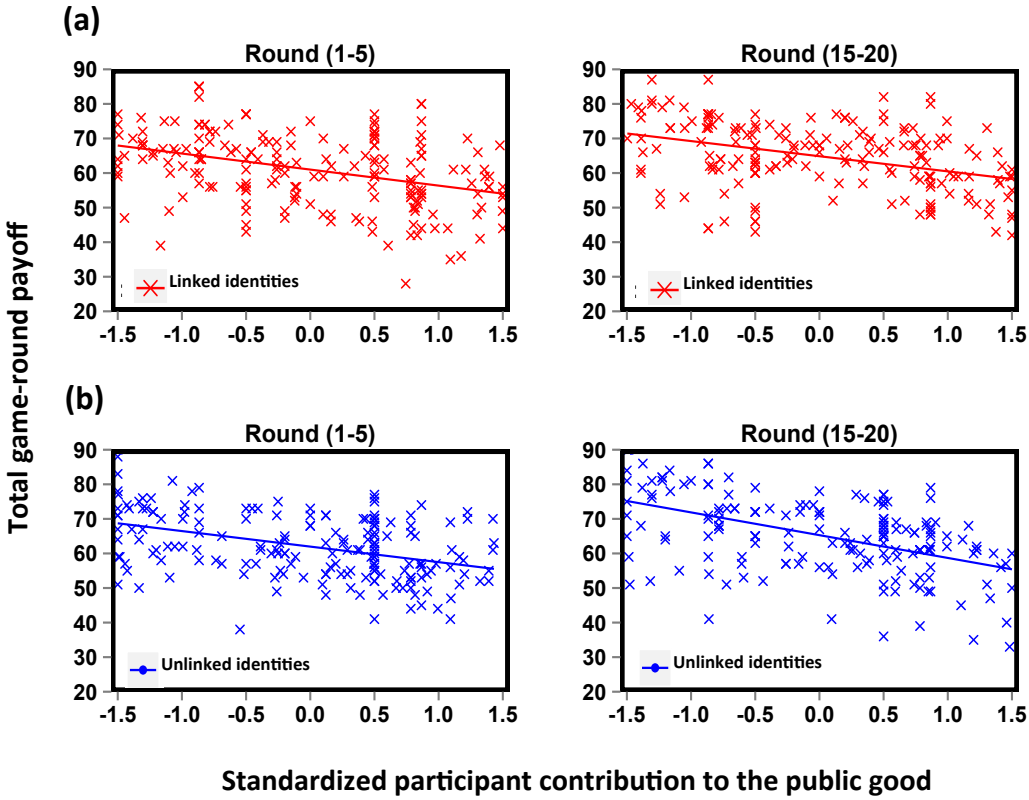


Figure 5.13. Scatterplots showing participants' standardized contributions and game-round payoff (points) during the first 5 rounds and final 5 rounds, for the (a) linked identity condition (red) and (b) unlinked identity condition (blue) (raw data values with linear lines of best fit)

5.5 Discussion

These findings, to my knowledge, provide the first evidence of a direct relationship between cooperative behaviour in collective action dilemmas and cultural transmission. In this study, individuals who were locally and presently (that is in the current round) cooperative received increased access to social information from group members, and performed better on a cumulative learning task, relative to free-riders, when cooperative reputations were revealed. Thus, this study provides evidence that human interactions are subject to reputation-based cultural learning processes (i.e.,

access to cultural information is moderated by cooperative reputation). Contrary to expectations, the increased risk to free riders of being excluded from cultural information did not stabilise cooperation relative to a condition that prevented individuals from using reputational information in cultural decision-making.

The most plausible reason that reputation-based costs failed to stabilise cooperation in this study is that participants perceived the overall game-round advantage of free riding: free riders still scored higher overall, relative to co-operators, despite the negative impact of reputational costs on cultural learning and/or the symbolic cost of being excluded from information by group members. This result contradicts a previous study arguing that participants do respond positively to symbolic sanctions in a cooperative dilemma – i.e., of anticipating written feedback from the receiver of their offer – increasing their offer as a result (Ellingsen & Johannesson, 2008). However, important procedural differences between the two studies could easily explain this finding; namely that the economic game used in Ellingsen and Johannesson's study (the ultimatum game) differed in important respects from the PGG used here (i.e., player offers could be all-out rejected leaving both players with nothing).

Nonetheless, the results presented here do not invalidate the hypothesis that cultural evolutionary processes do play a role in bootstrapping large-scale cooperation. For instance, it is conceivable that cultural information exchange can also occur in circumstances that do incur considerable costs to the receiver, should the information be denied. A hypothetical situation might involve excluding individuals from information that would allow them to partake and benefit from a group food hunt; or of denying individuals information about an impending danger. Under the same dynamics observed in this experiment, it is expected that increasing the cost of

being denied information would serve to stabilise or increase cooperation. Future studies should assess the effect that varying the cost/benefit of social information has on game dynamics and cooperation. Other interesting avenues of research might involve manipulating the group-level benefits of cooperation, to address questions such as: Can cooperative behaviour still act as a symbolic marker of reputation in cultural processes when the benefits of cooperation are negligible?

Further support of a role for cooperative reputation signalling in cultural learning comes from the observation that, after controlling for differences between conditions in absolute levels of information sharing, participants in this study were more willing to share high-quality information (i.e., high scoring monster teams) when they knew the reputation of the information recipient than when they didn't. Thus, participants who could see the reputations of group members were more selective overall about whom they shared with (preferentially sharing with co-operators), but were also relatively more willing to share high-scoring information than participants who did not have access to the reputations of their information recipients.

To my knowledge, these results provide the first evidence that humans not only *use* social information selectively (Boyd & Richerson, 1985; Morgan et al., 2012), but also *share* information selectively too; fine-tuning their decisions according to a combination of factors pertaining to both the quality of the information available to share, and some evaluation (here of the cooperative reputation) of the information recipient. This finding could at least partly explain the observation that performance in the learning game was similar between experimental conditions, despite higher levels of information transfer in the *unlinked identity* condition. It could also explain the finding that information sharing was positively related to learning

game performance when the recipients' reputations were visible, but not when they were hidden.

Previous theoretical investigations imply that individuals' reliance on social learning should be flexible, and that social learning might only feature prominently in learning decisions when individual learning has proven too costly or failed (Boyd & Richerson, 1995; Enquist et al., 2007; Laland & Kendal, 2003; Laland, 2004; although see Rendell et al., 2010a for a contrary view).The structure of the current learning game – namely the stable environment and score distribution experienced throughout; the constant availability of participants' previous best scoring team; and the risk free, individual learning practice period – might therefore have reduced participants' feeling of dependency or reliance on social learning and contributed to the unexpected null effect of reputation-based learning on cooperation. Future work should also examine the interaction between social learning dependence and the reputation-based costs of free riding on cooperation. Interestingly, a previous study suggests that individuals who favour increased reliance on social learning, also self-report as being more collectivist (Toelch et al., 2014).

The results presented here complement previous empirical work that indicates a relationship between pro-social behaviour and complex culture. For example, Dean et al. (2012) present evidence indicative of a link between altruistic information transfer and cumulative knowledge gain. Similarly, Hill et al. (2011, 2014) have argued that a unique social structure, which enhances large-scale, cooperative interactions between non-relatives, has scaffolded and coevolved with cultural complexity in human society. The results of this study extend these previous findings by providing mechanistic evidence of a reciprocal relationship between cooperative reputation and cultural information transfer. Information provision within groups or

societies of learners is not indiscriminate; rather, the cooperative reputation of the information recipient might be just as important as the cooperativeness of the information donor, or indeed network of donors, in many situations.

Of course, there are many situations where cultural information exchange itself might constitute a cooperative or information-sharing dilemma (Cabrera & Cabrera, 2002; Rendell et al., 2011b): individual learners are sometimes faced with the conundrum of contributing useful information for the collective benefit of the group, or withholding that same information to maintain a competitive edge over others. Likewise, teaching and language – key instruments of human cultural exchange (Morgan et al., 2015a; Tomasello, 1999) – are themselves considered major cooperative feats in their own right (Sterelny, 2003). Indeed, many of the cultural skills that emanate from cumulative learning (e.g., learning to make stone tools) require sustained and repeated interactions between learners and their expert teachers (Stout, 2002); plausibly, the learner's reputation is key in maintaining this process across kinship boundaries in the absence of formal institutions. Thus, reputation-based social transmission is likely ubiquitous and intrinsic throughout the complete set of processes and interactions that together constitute our cultural matrix.

An important final consideration in light of these results is of course the extent to which it is possible for individuals to control the flow of cultural information within and between group members. Whilst it is certainly the case that some forms of cultural information are not monopolisable, it is equally the case that many forms of cultural learning are dependent on active demonstration and communication. Moreover, even cultural information that is freely available for all to learn may still hinge on individuals maintaining close and cooperative spatial relations, to enable social transmission to occur (Hill et al., 2014; Rand et al. 2011).

In sum, these data provide evidence that cultural transmission processes are governed not only by selective learning biases present in the learner, but also biases determining when knowledgeable individuals are willing to transmit (or share) valuable information. In this study, individuals were more likely to actively share information with individuals they knew had contributed cooperatively to a public good. Moreover, when reputation-based decision-making was possible, public cooperation showed a positive relationship with group and individual performance in a cultural learning game.

It still remains to be proven whether the reputation-based costs inherent in cultural learning processes can serve to stabilise or increase cooperative behaviour. The benefits conferred by free riding in this study were not offset by the increased risk of being denied access to cultural information. However, given that reputation-based sharing is likely pervasive throughout the multiple forms of information exchange that scaffold human cultural behaviour, it is highly conceivable that many exchanges do risk considerable costs, and therefore do discourage free riders.

Within the cultural dynamics described here, it is easy to imagine a coevolutionary process between increasingly cooperative societies and increasingly powerful culture, with the roles of sanctioning free riders and distributing cultural knowledge eventually being supported by the emergence of cultural norms and formal institutions (Chudek & Henrich, 2011). It is further conceivable that this relationship would lead some cultural groups to triumph over others during times of inter-group conflict and competition, should enough variation exist between groups, further amplifying the evolutionary process (e.g., Bell et al., 2009; Bowles, 2006; Boyd & Richerson, 1985; Henrich, 2004b).

Chapter 6

General Discussion

In the previous chapters, I presented data examining the role of learning mechanisms in high-fidelity social transmission, the role of learning biases in 'over-imitation', and a series of experiments examining the interaction between human social learning and cooperation. In this chapter, I will first provide a recap of these findings. I will then consider some broader implications of the work presented, and some areas where future research is required, including two follow-up projects that are already underway.

6.1. Summary of findings

The focus of **Chapters 2** and **3** sits broadly within the topic of high-fidelity social learning, and is relevant to discussions concerning the importance of high-fidelity transmission in cultural ratcheting. The ability of individuals to engage in accurate or high-fidelity social transmission is considered necessary to both initiate and fuel the cultural ratcheting process (e.g., Lewis & Laland, 2012; Tomasello, 1999). The psychological and behavioural mechanisms thought to underpin high-fidelity learning (e.g., imitation) have also been proposed, and investigated across a range of species (frequently using the two-action paradigm; Dawson & Foss, 1965). While the focus has been on instrumental or technological cultural ratcheting, attempts to link the proposed mechanisms underpinning fidelity to increases in the faithful transmission of instrumental goals have been neglected, and the few studies attempting to examine this relationship have returned inconsistent results (Caldwell & Millen, 2009; Fredman & Whiten, 2008; Morgan et al., 2015b)

In **Chapter 2**, I contributed to answering this question using results obtained from a meta-analysis conducted across primate social learning studies that had been conducted using the two-action paradigm. This analysis showed that matching the demonstrated process of task solution (i.e., imitative or emulative learning – mechanisms associated with high-fidelity transmission) is not necessary for successful transmission of instrumental goals (i.e., reward retrieval) when the puzzle is simple, but becomes increasingly necessary as task complexity increases. This result remained whether human subjects were included or excluded from the analysis, and provides support for the view that mechanisms resulting in greater behavioural fidelity to the demonstrated process (i.e., imitation/emulation) are required, and have been selected for, in the evolution of instrumental culture.

'Over-imitation' describes the propensity of humans to blanket copy even actions that are manifestly irrelevant to achieving an instrumental outcome. It has been proposed that this puzzling phenomenon might represent a special case of high-fidelity learning in humans that operates to ensure the transmission of causally opaque artefact functions, despite sometimes leading to causally inefficient behaviours being copied (Horner & Whiten, 2005; Lyons et al., 2007). However, in **Chapter 3** I presented evidence that questions this interpretation, suggesting instead that children might copy or discard the demonstration of irrelevant actions in a manner that is compatible with the operation of selective and adaptive learning heuristics. Adopting a cultural evolutionary approach, I found that children copied the majority behaviour selectively, showing a bias for the majority's solution in all instances where it represented behavioural efficiency, but not when it involved copying causally irrelevant actions. Children's copying of the irrelevant action also showed sensitivity to the frequency of demonstrators who performed it, in a manner consistent with

social rather than causal explanations of 'over-imitation' (e.g., affiliative or normative motivations, e.g., Nielsen & Blank, 2011; Kenward, 2012). Using my results, I argued that children's tendency to copy irrelevant actions might be more compatible with the operation of a learning heuristic that produces 'optimal-' rather than 'over-' imitation.

In **Chapters 4** and **5**, the focus of the thesis shifted towards the relationship between cultural transmission and cooperation, and the role of this interaction in sustaining complex culture, using experimental investigations with adult human participants. In the General Introduction, I argued that enhanced cooperativeness and enhanced cultural complexity had likely coevolved in humans, in a mutually reinforcing, reciprocal manner, but that this relationship was currently under-investigated experimentally.

In **Chapter 4** I attempted to test the hypothesis that increases in cooperative rapport can lead to increases in imitative behaviour (of gestures and mannerisms) between interaction partners (van Baaren et al., 2009; Heyes, 2012a), using a series of three experiments. There are already a number of empirical studies offering support for the reciprocate of this relationship – that leading from imitation to increased cooperation (e.g., Carpenter et al., 2013; van Baaren et al., 2004) – and it has been argued that these two processes are engaged in a mutually reinforcing circle of cooperation and imitation, which might act as a springboard for the emergence of more complex forms of cultural learning (Heyes, 2012a).

In the first experiment, imitation was positively associated with a measure of the quality of participants' cooperative interaction (i.e., cooperative task performance). Moreover, individuals in a control condition, who performed the task alone, later imitated their interaction partner more if they themselves had performed poorly on the task, consistent with the hypothesis that experiencing failure increases

individuals' desire to affiliate with others through imitation (Lakin & Chartrand, 2003). However, this initial finding was not replicable in later experiments. The experiments also failed to detect the expected increase in imitative behaviour among participants who cooperated, relative to the control group, despite a series of incremental experimental modifications designed to enhance participants' cooperative experience. This finding contrasts with a previous study showing an increase in imitative behaviour in participants primed with cooperative words (Leighton et al., 2010). It is possible that the relationship between cooperation and imitation is more nuanced than the behavioural interactions permitted in my experiments, preventing the detection of replicable and expected results.

In **Chapter 5**, I presented data investigating the role that cultural transmission might play in helping to sustain cooperation. Using a micro-society approach (Mesoudi & Whiten, 2008), involving groups of participants playing a networked computer game, I examined whether free riders in a public goods game dilemma were more likely to be excluded from receiving potentially useful social information in a separate but linked learning game. Moreover, I was interested in whether the threat of reputation-based exclusion from social information could serve to discourage free riding and increase cooperation, relative to a control condition where reputation-based social transmission was not possible. I reasoned that humans' extreme dependence on social information might serve to encourage the maintenance of cooperative reputations, if access to beneficial social information was at stake, and could potentially provide a cheaper mechanism for sustaining cooperation than costly punishment or reward.

The results of this experiment showed that participants did exhibit biases regarding *whom* they were willing to share social information with, preferentially

sharing with co-operators and withholding information from free riders, as predicted. There also appeared to be evidence suggesting that individuals were relatively more willing to share high-value information with recipients whose reputations were visible, relative to when recipients' reputations were hidden. However, contrary to expectations, the operation of reputation-based access to social information in this experiment did not increase or stabilise cooperative behaviour, relative to the control condition. A closer examination of game-round payoffs offered an explanation for these contradictory findings, suggesting that the cost associated with exclusion from social information was not large enough to offset the benefit of free riding within the game set up employed, and that participants likely perceived this. The symbolic cost of experiencing targeted exclusion from information was not enough by itself to deter free riders. However, the fact that reputation-based access to information still appeared to be operating is indicative that this mechanism could still plausibly function to sustain cooperation if exclusion from social information carried more significant social or survival costs; a situation that is conceivable both over the course of modern human history through to contemporary times.

6.2. Broader considerations

High-fidelity social transmission

The work presented in both Chapters 2 and 3 is of relevance to discussions concerning the evolution of high-fidelity social transmission, and its role in cultural complexity. Chapter 2, in addition to providing support for a relationship between learning mechanisms and instrumental learning success, highlights again the difficulty of differentiating between imitative and emulative learning mechanisms, and again raises the question of whether this distinction is useful in discussions of instrumental social learning that utilise current paradigms.

My own opinion is that the imitation-emulation dichotomy, as it currently stands in assessments of instrumental social learning, is not especially useful, leads to conflation and confusions in terminology, and likely serves to muddy our understanding of where the differences between human and non-human copying truly lie. The observation that this dichotomy is being conflated by even those authors who stress its importance (e.g., Tennie et al., 2009), further indicates that currently we lack the evidence, and also the necessary experimental paradigms, to demarcate human versus nonhuman learning on the grounds of an imitation/emulation learning dichotomy. Chapter 2 indicates that both human and non-human primates are able to exhibit fidelity to a demonstrated process of task solution, and that this fidelity becomes increasingly necessary for the transmission of instrumental outcomes as task complexity increases. This finding supports the common assumption that mechanisms such as imitation/emulation afford higher fidelity transmission of cultural learning goals and outcomes than mechanisms such as stimulus/local enhancement. The use of conditions offering reduced social information, such as ghost conditions, appear to

result in a reduction in learning performance for both human and nonhuman species, yet these manipulations have been criticised for unrealistically distorting the natural process of social learning (Byrne, 2002), and still fail to provide an accurate assessment of the relative contributions of different learning mechanisms when learners have all options available (i.e., during a complete demonstration). Our current experimental paradigms lack the empirical resolution to make the imitation-emulation distinction accurately.

Indeed, it would appear that our knowledge of the catalogued learning mechanisms, including of the psychological and neural mechanisms on which they hinge, is still rather limited (Heyes, 2012b; Hoppitt & Laland, 2008). Even within the category of imitation – the most studied and discussed type of copying – there is still uncertainty regarding whether the different types of *simple* imitation (e.g., behavioural mimicry, response facilitation, automatic imitation) arise from the same or different neural circuits, and of exactly where *simple* imitation ends and *complex* imitation begins (Heyes, 2011). Elsewhere, the presumed cognitive superiority of imitative relative to emulative learning has been challenged, along with the assumption that the two mechanisms result from different underlying cognitive processes (Hoppitt & Laland, 2008). Recreating a pattern of observed object movements, or recreating an end product, could conceivably be as cognitively complex, or even more cognitively complex, than matching a demonstrator's action topography. The point is that, while the distinction between emulation and imitation might be an important one, this is still very much an open empirical question. My own hunch is that humans have evolved enhanced abilities in both domains, and that the relative importance of imitation and emulation as mechanisms of social learning (both in instrumental and in social/ritual contexts), varies systematically dependent upon the details of the learning goal. Until

the development of experimental paradigms that allow the relative role of each mechanism to be deciphered, the current dogma surrounding the imitation/emulation dichotomy (e.g., Tennie et al., 2009) is unlikely to be a useful one.

Of course, increasingly the role of other socio-cognitive abilities – such as intentional teaching through verbal instruction and pro-social interactions – considered particularly enhanced in humans (Tomasello, 1999), are coming to the fore in explanations of uniquely human cultural affairs, offering pathways to high-fidelity transmission and cultural ratcheting that supplement and supersede those offered by imitation and emulation alone (e.g., Dean et al., 2012; Morgan et al., 2015a; Zwirner & Thornton, 2015). Future research must seek to elucidate further the complete suite of socio-cognitive abilities that operate to furnish our species with culture (Tomasello, 1999).

It is proposed that 'over-imitation' in humans, or perhaps more accurately 'over-emulation' or 'over-copying', functions to ensure the high-fidelity transmission of opaque artefact functions and behaviours that frequently characterise human culture (e.g., Lyons et al., 2011). According to this hypothesis, humans display an evolved tendency towards automatic or blanket copying, which usually functions to ensure that all functional properties of a behavioural sequence are adopted, but can also occasionally result in manifestly irrelevant behaviour being copied. The evidence presented in Chapter 3 challenges this view, suggesting that causally irrelevant action copying is selective, and instead might function to serve social (e.g., affiliative/conventional) rather than instrumental goals, in line with an alternative body of literature on this topic (e.g., Keupp et al., 2013; Nielsen & Blank, 2011; Over & Carpenter, 2012). There is still clearly more work required in unravelling exactly why humans copy causally irrelevant actions; the suggestion that multiple

explanations exist, which are underpinned by different psychological learning mechanisms and vary in relevance according to learning context, is certainly plausible (Kenward, 2012).

However, claims that irrelevant action copying results from a misguided tendency to blanket copy actions with high-fidelity during causal learning – i.e., instances that truly depict 'over-imitation' – need to be distinguished from cases where the irrelevant action has incorporated a socially relevant function or significance, which I earlier argued is better depicted as 'optimal-' rather than 'over-' imitation. There is also a need to clarify and better quantify the so-called behavioural costs incurred when individuals copy irrelevant actions; which might be relatively low or negligible despite being what has elevated 'over-imitation' to its position as a topic of high interest. While previously it was argued that over-imitation persists despite substantial costs of performing the irrelevant action (Lyons et al., 2011), a more recent study found that children's propensity to copy irrelevant actions decreased as the cost incurred from doing so became increasingly substantial (Keupp et al., 2016). A further issue that confuses interpretation arises when the causally irrelevant actions offered in claims of 'over-imitation' are not in fact causally transparent or manifestly irrelevant – here evidence of causally irrelevant action copying is better described as part of the normal causal learning process (Kenward et al., 2011). There is clearly a requirement in research on 'over-imitation' for a more detailed outline on what exactly 'over-imitation' *is* and *isn't*. However, with the above considerations in mind, and the additional finding that selective learning biases likely filter irrelevant actions adaptively (Chapter 3), my guess is that instances of blanket copying causally irrelevant actions, for causally functional reasons (i.e., truly over-imitating), are likely few and far between, both in the 'real world' and in the laboratory.

Social learning strategies

A recent series of articles has taken issue with what the authors have interpreted as the depiction of cognitively fixed and inflexible social learning heuristics in the current literature (e.g., Heyes & Pearce, 2015; Heyes, 2016). While this depiction of the current literature is itself a moot point, the results presented in Chapter 3 argue to the contrary: that social learning heuristics or strategies are in fact highly flexible, and subject to substantial alteration in learning contexts that involve irrelevant information (which also plausibly extends to other forms of disadvantageous or context-irrelevant information). Indeed, a recent review argues that, rather than assuming species-typical learning heuristics, a growing body of evidence implies much individual-level variability in social learning, which plausibly results from a range of influencing factors; including heritable differences in social learning, but also differences resulting from developmental, individual learning experiences, and, in humans, cultural learning experiences (Mesoudi et al., 2016).

Selective or strategic copying biases have been identified across a range of different species (See General Introduction, *Section 1.4)*. In humans, it is believed they play a fundamental role in cultural evolutionary processes at the population level, such as in ensuring the stability of culture and in guiding learning towards the local optima (Boyd & Richerson, 1985; Chudek & Henrich, 2011). The challenge now facing researchers is to elucidate how selective copying in humans has evolved to be different from that observed in other animals, in ways that foster increasing cultural complexity. Recent simulations undertaken on this issue have indicated that increases in highly selective and refined copying strategies, along with a greater reliance on social information more generally, are two such candidate factors (Rendell et al., 2010a).

One prime candidate for how humans might achieve more refined and selective copying heuristics is through the operation of human-specific, metacognitive decision-making processes in social learning (Heyes, 2016); an understudied area ripe for future research. Another obvious area of consideration is in the uniquely human capacity for culturally honed and transmitted rules of social learning (Mesoudi et al., 2015), including the ways in which our specialised 'norm psychology' fashions social transmission (Chudek & Henrich, 2011).

In addition to understanding the processes by which individuals choose whom to copy from, there may also be a requirement to understand the operation of biases that influence to whom one should award access to cultural information. Chapter 5 identified the operation of reputation-based biases in information sharing: individuals preferentially shared information with co-operators and withheld information from free riders (as determined by behaviour in a public goods game). It is plausible that a number of different biases might operate regarding to whom one should supply cultural information; in addition to the reputation-based biases identified in Chapter 5, kin- and prestige-based information-sharing biases would seem other likely candidates. Further investigation of these potential biases, and consideration of their likely impact on cultural evolutionary processes, are required in future.

Cooperation and culture

The work presented in Chapter 5, which utilised a public goods game linked to a separate learning game, is currently undergoing extension and follow up (see Section 6.3, below). Of relevance to the current and follow up study is a recent series of experiments which question the use of laboratory-based economic games in investigations of cooperative behaviour (Burton-Chellew & West, 2013; Burton-

chellew et al., 2016). Specifically, Burton-Chellew and colleagues take issue with the view that participant behaviour in economic games – namely the tendency of participants to initially cooperate at much higher levels than is individually or economically rational – indicates the evolution of strong pro-social tendencies in our species, arguing instead that this behaviour more likely results from an initial misunderstanding of the game rules about how to maximise individual payoffs. As the primary purpose of my investigation was to examine the occurrence of reputation-based access to social information and its role in stabilising cooperation, the debate about whether humans can correctly be characterised as 'irrationally' pro-social is of secondary importance to my question. However, the issues raised do illustrate the need to assess individuals' understanding of game dynamics when interpreting the patterns of decision-making that emerge from laboratory experiments, and when using these patterns to make inferences about behaviour in the 'real world'.

On a related note, the field of cultural evolutionary studies has been criticised for being over-reliant on theoretical modelling approaches that are frequently not validated in real-world settings (Mace, 2014). I think this same concern, or caution, is also relevant to laboratory-based micro-society approaches to the study of cultural evolution, which I adopted in Chapter 5. While the micro-society approach offers an extremely valuable and logistically feasible tool for investigating the mechanisms and processes that might underpin the evolution of large-scale behavioural patterns (Mesoudi & Whiten, 2008), the results that emerge from these microcosms, and that often appear in high profile journals (e.g., Derex et al., 2013; Derex & Boyd, 2016), might frequently depend, as is also the case in theoretical investigations, on the way in which we have created or parameterised our model of the 'micro' world. Taking the example of my own study (Chapter 5), if it is shown that reputation-based access to

social information can stabilise or increase cooperation, but only when experimental parameters are such that the value of receiving social information, relative to free riding, is high (see Chapter 5), we can conclude that this mechanism *might* have been important during the coevolution of cultural complexity and large-scale cooperation. However, the question we are then left with is: Is the value of receiving social information high enough, relative to free riding, to sustain this dynamic in the 'real world'? Answering this type of question is not easy, but recent attempts to shift the methods of economic and social learning experiments into 'real world' (or field) settings, with 'real world' commodities, are a step in the right direction (e.g., Lamba & Mace, 2011; Lamba, 2014).

6.3. Further work

Work already underway

Two studies that follow from the work presented in previous chapters are already underway. The first involves an extension of the work presented on 'over-imitation' in Chapter 3. The previous protocol involved presenting the task using outcome-oriented language, to emphasise the instrumental end-goal (Clegg & Legare, 2015), finding that children did not show majority-biased copying if the majority performed an irrelevant action. However, in the current follow-up study, children's tendency to copy irrelevant actions performed by a majority is compared across conditions that emphasise either an instrumental *or* a normative stance. Majority-biased or conformist copying can serve to guide both instrumentally driven and socially motivated learning goals, described in social psychology as *informational* and *normative* conformity

respectively (Morgan & Laland, 2012). In line with the idea that children might be better viewed as 'optimal-imitators', it is expected that when the majority behaviour is presented using normative cues, children will show majority-biased copying of the irrelevant action, as this now incorporates socially relevant functional properties, but will continue not to show majority-biased copying when the goal is instrumental.

The second follow-up study, involves an extension of the work presented in Chapter 5. Chapter 5 found that, although participants' access to social information appeared to be mediated by their reputation in a linked public goods, the cost to free riders of being denied social information was not sufficient to discourage free riding and stabilise cooperation. In the follow-up study, the original social learning game (i.e., the *Monster League Game*) has been replaced by a different custom-made computer game – the *Food Hunt Game* – which involves players 'hunting' for 'food items' and 'tools' upon a virtual memory board. In this game, players are permitted to sample game board locations before making performance-related game decisions, and are awarded points for locating complementary 'food' and 'tool' items, but can also lose points if they uncover 'dangerous' items. Players are permitted to share information about 'food', 'tool' and 'dangerous' game locations, as per the previous experimental design. Critically, the new game design offers the option of varying the costs and benefits of receiving social information (e.g., the possibility of losing a large number of points when sampling dangerous locations results in social information about these locations becoming particularly valuable), which is expected to permit a more detailed and refined assessment of the potential interaction between cooperative behaviour and cultural transmission. It is predicted that reputation-based access to social information will act to stabilise cooperation when the value of receiving social

information outweighs the benefits of free riding, but not when the value of social information is relatively low.

The need for more fine-grained studies of learning and behaviour

A recurrent theme throughout this thesis has been to contrast the emphasis placed on the specific mechanisms underpinning learning (e.g., imitation/emulation) with the inadequacy of current experimental methods to accurately categorise the learning process. More refined and detailed studies are clearly required to unravel and document the precise learning mechanisms and processes that underpin social transmission, and to determine more accurately how these might differ between species. One approach that might prove useful in future is the implementation of sophisticated eye tracking technology, which would allow researchers to track and document exactly where, and how much, individuals focus their attention during social learning – e.g., on object movements (emulation) or body movements (imitation) – across a range of different learning objectives. Such methods could also be expanded, at least in humans, to include a more holistic consideration of the learning dynamic that occurs between individuals and their cultural models, incorporating a detailed evaluation of the role played by behaviours such teaching, verbal communication and joint attention, in supplementing social learning mechanisms during cultural transmission. While the development and implementation of such an experimental framework is likely to be complex, it might prove essential if we wish to develop a deeper understanding of the suite of learning processes that underpin human culture.

6.4. Conclusions

The work presented in this thesis contributes to two main themes that are of relevance to understanding the evolution of human cultural uniqueness. The first theme involves the evolution of high-fidelity cultural transmission, while the second theme concerns the interaction between cultural transmission and cooperation. This thesis supports previous, but largely untested, claims that learning mechanisms associated with increased fidelity to a cultural model's behaviour (i.e., imitation/emulation) are required to learn and transmit complex instrumental cultural end-goals successfully. However, it also questions the current dogma of casting humans as 'imitators' and nonhumans as 'emulators', in light of the current evidence available in social learning research. This thesis also questions the depiction of human beings as 'over-imitators', and the suggestion that 'over-imitation' represents an adaptation for the high-fidelity transmission of causally opaque cultural behaviours. Instead, it finds evidence that the demonstration of causally irrelevant actions might substantially alter the operation of adaptive learning biases, in ways that likely reduce or eliminate 'over-imitation', and show sensitivity to imitation's social, rather than causal, functions. In later chapters, the thesis finds only limited support for a positive relationship between cooperative interactions and automatic imitation, which is hypothesised to function in maintaining cultural group cohesion; a result that might reflect nuances in this relationship that were missing from the experimental design. Finally, the thesis finds support for the hypothesis that access to cultural information is mediated by an individual's cooperative reputation, and argues that this interaction could plausibly serve to help scaffold the evolution of both large-scale cooperation and complex culture, if being denied access to cultural information frequently carries considerable costs. The

overarching message presented throughout this thesis is that understanding human cultural uniqueness hinges upon understanding a complex of individual- and population-level processes, from individual learning mechanisms to large-scale cooperation, which researchers are only just beginning to unravel.

References

Akins, C.K. & Zentall, T.R. (1998). Imitation in Japanese quail: The role of reinforcement of demonstrator responding. *Psychonomic Bulletin & Review*. 5 (4). pp. 694–697.

Akins, C.K. & Zentall, T.R. (1996). Imitative learning in male Japanese quail (Coturnix japonica) using the two-action method. *Journal of Comparative Psychology*. 110 (3). pp. 316–20.

Allen, J., Weinrich, M., Hoppitt, W. & Rendell, L. (2013). Network-based diffusion analysis reveals cultural transmission of lobtail feeding in humpback whales. *Science*. 340 (6131). pp. 485–8.

Apicella, C.L., Marlowe, F.W., Fowler, J.H. & Christakis, N. a. (2012). Social networks and cooperation in hunter-gatherers. *Nature*. 481 (7382). pp. 497–501.

Aplin, L.M., Farine, D.R., Morand-Ferron, J., Cockburn, A., Thornton, A. & Sheldon, B.C. (2015a). Counting conformity: evaluating the units of information in frequency-dependent social learning. *Animal Behaviour*. 110. pp. e5–e8.

Aplin, L.M., Farine, D.R., Morand-Ferron, J., Cockburn, A., Thornton, A. & Sheldon, B.C. (2015b). Experimentally induced innovations lead to persistent culture via conformity in wild birds. *Nature*. 518 (7540). pp. 538–541.

van Baaren, R., Janssen, L., Chartrand, T.L. & Dijksterhuis, A. (2009). Where is the love? The social aspects of mimicry. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 364 (1528). pp. 2381–9.

van Baaren, R.B., Fockenberg, D., Holand, R., Janssen, L. & van Knippenberg, A. (2006). the Moody Chameleon: the Effect of Mood on Non-Conscious Mimicry. *Social Cognition*. 24 (4). pp. 426–437.

van Baaren, R.B., Holland, R.W., Kawakami, K. & van Knippenberg, A. (2004). Mimicry and prosocial behavior. *Psychological Science*. 15 (1). pp. 71–4.

Bartoń, K. (2015). MuMIn: multi-model inference: Model selection and model averaging based on information criteria (AICc and alike), R package version1.15.1. Appendix material

Bates, D., Maechler, M., Bolker, B., & Walker S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7, http://CRAN.R-project.org/package=lme4

Beck, S.R., Apperly, I.A., Chappell, J., Guthrie, C. & Cutting, N. (2011). Making tools isn't child's play. *Cognition*. 119 (2). pp. 301–6.

Bell, A. V, Richerson, P.J. & McElreath, R. (2009). Culture rather than genes

provides greater scope for the evolution of large-scale human prosociality. *Proceedings of the National Academy of Sciences of the United States of America*. 106 (42). pp. 17671–4.

Bell, C. & Sternberg, E. (2001). Emotional selection in memes: the case of urban legends. *Journal of Personality and Social Psychology*. 81 (6). pp. 1028–41.

Belot, M., Crawford, V.P. & Heyes, C. (2013). Players of Matching Pennies automatically imitate opponents' gestures against strong incentives. *Proceedings of the National Academy of Sciences of the United States of America*. 110 (8). pp. 2763–2768.

van Bergen, Y., Coolen, I. & Laland, K.N. (2004). Nine-spined sticklebacks exploit the most reliable source when public and private information conflict. *Proceedings of the Royal Society B: Biological Sciences*. 271 (1542). pp. 957–62.

Birch, S.A.J., Vauthier, S.A. & Bloom, P. (2008). Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition*. 107 (3). pp. 1018–34.

Bird, G., Leighton, J., Press, C. & Heyes, C. (2007). Intact automatic imitation of human and robot actions in autism spectrum disorders. *Proceedings of the Royal Society B: Biological Sciences*. 274 (1628). pp. 3027–31.

Biro, D., Haslam, M. & Rutz, C. (2013). Tool use as adaptation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 368 (1630). pp. 20120408.

Biro, D., Inoue-Nakamura, N., Tonooka, R., Yamakoshi, G., Sousa, C. & Matsuzawa, T. (2003). Cultural innovation and transmission of tool use in wild chimpanzees: evidence from field experiments. *Animal Cognition*. 6 (4). pp. 213–23.

Blomberg, S.P., Garland, T. & Ives, A.R. (2003).Testing for phylogentic signal in comparative data: Behavioural traits are more labile. *Evolution*. 57 (4). pp. 717–745.

Bloomberg, S. & Garland, T. (2002). Tempo and mode in evolution: Phylogenetic inertia, adaptation and comparative methods. *Journal of Evolutionary Biology*. 15. pp. 899–910.

Boesch, C. (2003). Is culture a golden barrier between human and chimpanzee? *Evolutionary Anthropology*. 12 (2). pp. 82–91.

Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H. & White, J.S.S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology and Evolution*. 24 (3). pp. 127–135.

Bonner, J. (1980). *The Evolution of Culture in Animals*. Princeton, New Jersey: Princeton University Press.

Bowles, S. (2006). Group competition, reproductive leveling, and the evolution of human altruism. *Science*. 314 (5805). pp. 1569–72.

Boyd, R. & Richerson, P.J. (1985). *Culture and the Evolutionary Process*. University of Chicago Press.

Boyd, R. & Richerson, P.J. (1996). Why culture is common, but cultural evolution is rare. In: *Proceedings - British Academy*. 1996, Oxford University Press, pp. 77–94.

Boyd, R. & Richerson, P.J. (1995). Why does culture increase human adaptability? *Ethology and Sociobiology*. 16 (2). pp. 125–143.

Brass, M., Bekkering, H., Wohlschläger, A. & Prinz, W. (2000). Compatibility between Observed and Executed Finger Movements: Comparing Symbolic, Spatial, and Imitative Cues. *Brain and Cognition*. 44 (2). pp. 124–143.

Brass, M. & Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*. 9 (10). pp. 489–95.

Bugnyar, T. & Huber, L. (1997). Push or pull: an experimental study on imitation in marmosets. *Animal Behaviour*. 54 (4). pp. 817–831.

Burkart, J., Kupferberg, A., Glasauer, S. & van Schaik, C. (2012). Even simple forms of social learning rely on intention attribution in marmoset monkeys (Callithrix jacchus). *Journal of Comparative Psychology*. 126 (2). pp. 129–38.

Burkart, J.M., Allon, O., Amici, F., Fichtel, C., Finkenwirth, C., Heschl, A., Huber, J., Isler, K., Kosonen, Z.K., Martins, E., Meulman, E.J., Richiger, R., Rueth, K., Spillmann, B., Wiesendanger, S. & van Schaik, C.P. (2014). The evolutionary origin of human hyper-cooperation. *Nature Communications*. 5. pp. 4747.

Burkart, J.M. & van Schaik, C.P. (2010). Cognitive consequences of cooperative breeding in primates? *Animal Cognition*. 13 (1). pp. 1–19.

Burnham, K. & Anderson, D. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theroretic Approach*. 2nd Ed. Springer.

Burton-Chellew, M. & West, S. (2013). Prosocial preferences do not explain human cooperation in public-goods games. *Proceedings of the National Academy of Sciences of the United States of America*. 110 (1). pp. 216–21.

Burton-chellew, M.N., El, C. & West, S.A. (2016). Conditional cooperation and confusion in public-goods experiments. *Proceedings of the National Academy of Sciences*. (Early online). pp. 1–6.

Byrne, R. (1995). *The thinking ape: Evolutionary origins of intelligence*. Oxford: Oxford University Press.

Byrne, R.W. (2002). Imitation of novel complex actions: What does the evidence from animals mean? *Advances in the Study of Behavior*. 31. pp. 77–105.

Byrne, R.W. & Russon, A.E. (1998). Learning by imitation: a hierarchical approach. *The Behavioral and Brain Sciences*. 21 (5). pp. 667–684.

Byrne, R.W. & Tomasello, M. (1995). Do rats ape? *Animal Behaviour*. 50 (5). pp. 1417–1420.

Cabrera, A. & Cabrera, E.F. (2002). Knowledge-Sharing Dilemmas. *Organization Studies*. 23 (5). pp. 687–710.

Caldwell, C. A & Millen, A.E. (2009). Social learning mechanisms and cumulative cultural evolution. Is imitation necessary? *Psychological Science*. 20 (12). pp. 1478–83.

Caldwell, C. A, Schillinger, K., Evans, C.L. & Hopper, L.M. (2012). End state copying by humans (Homo sapiens): implications for a comparative perspective on cumulative culture. *Journal of Comparative Psychology*. 126 (2). pp. 161–9.

Caldwell, C. a & Whiten, A. (2002). Evolutionary perspectives on imitation: is a comparative psychology of social learning possible? *Animal Cognition*. 5 (4). pp. 193–208.

Caldwell, C. A & Whiten, A. (2004). Testing for social learning and imitation in common marmosets, Callithrix jacchus, using an artificial fruit. *Animal Cognition*. 7 (2). pp. 77–85.

Caldwell, C.A. & Millen, A.E. (2010). Conservatism in laboratory microsocieties: unpredictable payoffs accentuate group-specific traditions. *Evolution and Human Behavior*. 31 (2). pp. 123–130.

Call, J., Carpenter, M. & Tomasello, M. (2005). Copying results and copying actions in the process of social learning: chimpanzees (Pan troglodytes) and human children (Homo sapiens). *Animal Cognition*. 8 (3). pp. 151–63.

Call, J. & Tomasello, M. (1994). The social learning of tool use by orangutans (Pongo pygmaeus). *Human Evolution*. 9 (4). pp. 297–313.

Cantor, M., Shoemaker, L.G., Cabral, R.B., Flores, C.O., Varga, M. & Whitehead, H. (2015). Multilevel animal societies can emerge from cultural transmission. *Nature Communications*. 6. pp. 8091.

Cantor, M. & Whitehead, H. (2013). The interplay between social networks and culture: theoretically and among whales and dolphins. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*. 368. pp. 1–10.

Carpenter, M., Akhtar, N. & Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development*. 21 (2). pp. 315–330.

Carpenter, M., Call, J. & Tomasello, M. (2005). Twelve- and 18-month-olds copy actions in terms of goals. *Developmental Science*. 8 (1). pp. F13–20.

Carpenter, M., Uebel, J. & Tomasello, M. (2013). Being mimicked increases

prosocial behavior in 18-month-old infants. *Child Development*. 84 (5). pp. 1511–1518.

Carvalho, P., Diniz-Filho, J.A.F. & Bini, L.M. (2006). Factors influencing changes in trait correlations across species after using phylogenetic independent contrasts. *Evolutionary Ecology*. 20 (6). pp. 591–602.

Cavalli-Sforza, L.L. & Feldman, M.W. (1981). Cultural transmission and evolution: a quantitative approach. *Monographs in Population Biology*. 16. pp. 1–388.

Chartrand, T. & Van Baaren, R. (2009). Human Mimicry. *Advances in Experimental Social Psychology*. 41 (08). pp. 219–274.

Chartrand, T.L. & Bargh, J. a (1999). The chameleon effect: the perception-behavior link and social interaction. *Journal of Personality and Social Psychology*. 76 (6). pp. 893–910.

Chudek, M. & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences*. 15 (5). pp. 218–226.

Claidière, N. & Sperber, D. (2010). Imitation explains the propagation, not the stability of animal culture. *Proceedings of the Royal Society B: Biological Sciences*. 277 (1681). pp. 651–9.

Clegg, J.M. & Legare, C.H. (2015). Instrumental and Conventional Interpretations of Behavior Are Associated With Distinct Outcomes in Early Childhood. *Child Development*. 87 (2). pp. 527–42.

Open Collaboration (2015). Estimating the reproducibility of psychological science. *Science*. 349 (6251). pp. 4716.

Cook, R., Bird, G., Lünser, G., Huck, S. & Heyes, C. (2012). Automatic imitation in a strategic context: players of rock-paper-scissors imitate opponents' gestures. *Proceedings of the Royal Society B: Biological Sciences*. 279 (1729). pp. 780–786.

Corriveau, K. & Harris, P.L. (2009). Choosing your informant: weighing familiarity and recent accuracy. *Developmental Science*. 12 (3). pp. 426–37.

Corriveau, K.H., Fusaro, M. & Harris, P.L. (2009). Going with the flow: preschoolers prefer nondissenters as informants. *Psychological Science*. 20 (3). pp. 372–7.

Corriveau, K.H. & Harris, P.L. (2010). Preschoolers (sometimes) defer to the majority in making simple perceptual judgments. *Developmental Psychology*. 46 (2). pp. 437–45.

Coultas, J.C. (2004a). When in Rome... An Evolutionary Perspective on Conformity. *Group Processes & Intergroup Relations*. 7 (4). pp. 317–331.

Coultas, J.C. (2004b). When in Rome... An Evolutionary Perspective on Conformity. *Group Processes & Intergroup Relations*. 7 (4). pp. 317–331.

Csibra, G. & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. *Processes of Change in Brain and Cognitive Development. Attention and Performance XX*. 21. pp. 249-274

Custance, D., Prato-Previde, E., Spiezio, C., Rigamonti, M.M. & Poli, M. (2006). Social learning in pig-tailed macaques (Macaca nemestrina) and adult humans (Homo sapiens) on a two-action artificial fruit. *Journal of Comparative Psychology*. 120 (3). pp. 303–13.

Custance, D., Whiten, A. & Fredman, T. (1999). Social learning of an artificial fruit task in capuchin monkeys (Cebus apella). *Journal of Comparative Psychology*. 113 (1). pp. 13–23.

Custance, D., Whiten, A., Sambrook, T. & Galdikas, B. (2001). Testing for social learning in the 'artificial fruit' processing of wildborn orangutans (Pongo pygmaeus), Tanjung Puting, Indonesia. *Animal Cognition*. 4 (3-4). pp. 305–313.

Dawson, B. V. & Foss, B.M. (1965). Observational learning in budgerigars. *Animal Behaviour*. 13 (4). pp. 470–474.

Dean, L.G., Kendal, R.L., Schapiro, S.J., Thierry, B. & Laland, K.N. (2012). Identification of the Social and Cognitive Processes Underlying Human Cumulative Culture. *Science*. 335 (6072). pp. 1114–1118.

Dean, L.G., Vale, G.L., Laland, K.N., Flynn, E. & Kendal, R.L. (2014). Human cumulative culture: a comparative perspective. *Biological Reviews of the Cambridge Philosophical Society*. 89 (2). pp. 284–301.

Derex, M., Beugin, M.P., Godelle, B. & Raymond, M. (2013). Experimental evidence for the influence of group size on cultural complexity. *Nature*. 503 (7476). pp. 389–391.

Derex, M. & Boyd, R. (2016). Partial connectivity increases cultural accumulation within groups. *Proceedings of the National Academy of Sciences*. 113 (11). pp. 201518798.

Derex, M., Godelle, B. & Raymond, M. (2012). Social learners require process information to outperform individual learners. *Evolution*. 67 (3). pp. 688–697.

Dillis, C., Humle, T. & Snowdon, C.T. (2009). Socially biased learning among adult cottontop tamarins (Saguinus oedipus). *American Journal of Primatology*. 72 (4). pp. 287-95.

Dindo, M., Stoinski, T. & Whiten, A. (2011). Observational learning in orangutan cultural transmission chains. *Biology Letters*. 7 (2). pp. 181–3.

Dindo, M., Thierry, B., de Waal, F.B.M. & Whiten, A. (2010). Conditional copying fidelity in capuchin monkeys (Cebus apella). *Journal of Comparative Psychology*. 124 (1). p.pp. 29–37.

Dindo, M., Thierry, B. & Whiten, A. (2008). Social diffusion of novel foraging methods in brown capuchin monkeys (Cebus apella). *Proceedings of the Royal*

Society B: Biological Sciences. 275 (1631). pp. 187–93.

Discamps, E. & Henshilwood, C.S. (2015). Intra-site variability in the Still Bay fauna at Blombos Cave : implications for explanatory models of the Middle Stone Age cultural and technological evolution. *PloS ONE*. 10 (12). pp. 1–21.

Dorrance, B.R. & Zentall, T.R. (2001). Imitative learning in Japanese quail (Coturnix japonica) depends on the motivational state of the observer quail at the time of observation. *Journal of Comparative Psychology*. 115 (1). pp. 62–7.

Einav, S. (2014). Does the majority always know best? Young children's flexible trust in majority opinion. *PloS ONE*. 9 (8). pp. e104585.

Ellingsen, T. & Johannesson, M. (2008). Anticipated verbal feedback induces altruistic behavior. *Evolution and Human Behavior*. 29 (2). pp. 100–105.

Enquist, M., Eriksson, K. & Ghirlanda, S. (2007). Critical Social Learning: A Solution to Rogers's Paradox of Nonadaptive Culture. *American Anthropologist*. 109 (4). pp. 727–734.

Enquist, M. & Ghirlanda, S. (2007). Evolution of social learning does not explain the origin of human cumulative culture. *Journal of Theoretical Biology*. 246 (1). pp. 129–35.

Enquist, M., Strimling, P., Eriksson, K., Laland, K. & Sjostrand, J. (2010). One cultural parent makes no culture. *Animal Behaviour*. 79 (6). pp. 1353–1362.

Eriksson, K., Enquist, M. & Ghirlanda, S. (2007). Critical points in current theory of conformist social learning. *Journal of Evolutionary Psychology*. 5 (1). pp. 67–87.

Fawcett, T.W., Skinner, A.M.J. & Goldsmith, A.R. (2002). A test of imitative learning in starlings using a two-action method with an enhanced ghost control. *Animal Behaviour*. 64 (4). pp. 547–556.

Fehr, E., Fischbacher, U. & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature*. 13 (1). pp. 1–25.

Felsenstein, J. (1985). Phylogenies and the Comparative Method. *The American Naturalist*. 125 (1). pp. 1–15.

Fitch, T, W. (2010). *The Evolution of Language*. Cambridge: Cambridge University Press.

Flynn, E. (2008). Investigating children as cultural magnets: do young children transmit redundant information along diffusion chains? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 363 (1509). pp. 3541–51.

Freckleton, R.P., Harvey, P.H. & Pagel, M. (2002). Phylogenetic analysis and comparative data: a test and review of evidence. *The American Naturalist*. 160 (6). pp. 712–26.

Fredman, T. & Whiten, A. (2008). Observational learning from tool using models by human-reared and mother-reared capuchin monkeys (Cebus apella). *Animal Cognition*. 11 (2). pp. 295–309.

Fusaro, M. & Harris, P.L. (2008). Children assess informant reliability using bystanders' non-verbal cues. *Developmental Science*. 11 (5). pp. 771–7.

Galef, B.G. (1988). Imitation in animals: History, definition, and interpretation of data from the psychological laboratory. In: T. R. Zentall & B. G. Galef (eds.). *Social Learning: Psychological and Biological Perspectives*. Hillsdale, NJ: Erlbaum, pp. 3–28.

Galef, B.G. (1990). Tradition in animals: field observations and laboratory analyses. In: M. Bekoff & D. Jamieson (eds.). *Interpretation and Explanation in the Study of Behavior*. Boulder: Westview Press, pp. 74–95.

Galef, B.G. (1992). The question of animal culture. *Human Nature*. 3 (2). pp. 157–78.

Garland, E.C., Goldizen, A.W., Rekdahl, M.L., Constantine, R., Garrigue, C., Hauser, N.D., Poole, M.M., Robbins, J. & Noad, M.J. (2011). Dynamic horizontal cultural transmission of humpback whale song at the ocean basin scale. *Current Biology*. 21 (8). pp. 687–691.

Gelman, A. & Rubin, D. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*. 7 (4). pp. 457–472.

Gergely, G., Bekkering, H. & Király, I. (2002). Rational imitation in preverbal infants. *Nature*. 415 (6873). pp. 755.

Gergely, G. & Csibra, G. (2006). Sylvia's recipe: The role of imitation and pedagogy in the transmission of cultural knowledge. In: N. Enfield & S. C. Levinson (eds.). *Roots of Human Sociality: Culture, Cognition and Interaction*. Berg Press, pp. 229–255.

Gintis, H., Bowles, S., Boyd, R. & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior*. 24 (3). pp. 153–172.

Giraldeau, L.-A., Valone, T.J. & Templeton, J.J. (2002). Potential disadvantages of using socially acquired information. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 357 (1427). pp. 1559–66.

Hadfield, J.D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software*. 33 (2). pp. 1–22.

Hastie, R. & Kameda, T. (2005). The robust beauty of majority rules in group decisions. *Psychological Review*. 112 (2). pp. 494–508.

Haun, D.B.M., Rekers, Y. & Tomasello, M. (2012). Majority-Biased Transmission in Chimpanzees and Human Children, but Not Orangutans. *Current Biology*. 22 (8). pp. 727–731.

Haun, D.B.M. & Tomasello, M. (2011). Conformity to peer pressure in preschool children. *Child Development*. 82 (6). pp. 1759–67.

Healy, S.D. & Rowe, C. (2007). A critique of comparative studies of brain size. *Proceedings of the Royal Society B: Biological Sciences*. 274 (1609). pp. 453–464.

Helfman, G.S. & Schultz, E.T. (1984). Social transmission of behavioural traditions in a coral reef fish. *Animal Behaviour*. 32 (2). pp. 379–384.

Henrich, J. (2004a). Demography and Cultural Evolution: How Adaptive Cultural Processes can Produce Maladaptive Losses: The Tasmanian Case. *American Antiquity*. 53 (1). pp. 1–35.

Henrich, J. (2004b). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization*. 53 (1). pp. 3–35.

Henrich, J. & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior*. 19 (4). pp. 215–241.

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 (06). pp. 795–815.

Henrich, J. & McElreath, R. (2003). The Evolution of Cultural Evolution. *Evolutionary Anthropology*. 12 (3). pp. 123–135.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D. & Ziker, J. (2006). Costly punishment across human societies. *Science*. 312 (5781). pp. 1767–70.

Herrmann, B., Thoni, C. & Gachter, S. (2008). Antisocial Punishment Across Societies. *Science*. 319 (5868). pp. 1362–1367.

Herrmann, E., Call, J., Hernàndez-Lloreda, M.V., Hare, B. & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science*. 317 (5843). pp. 1360–6.

Herrmann, P. a, Legare, C.H., Harris, P.L. & Whitehouse, H. (2013). Stick to the script: The effect of witnessing multiple actors on children's imitation. *Cognition*. 129 (3). pp. 536–43.

Heyes, C. (2011). Automatic Imitation. *Psychological Bulletin*. 137 (3). pp. 463–83.

Heyes, C. (2009). Evolution, development and intentional control of imitation. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*. 364 (1528). pp. 2293–8.

Heyes, C. (1993). Imitation, culture and cognition. *Animal Behaviour*. 46 (5). pp. 999–1010.

Heyes, C. (2012a). What can imitation do for cooperation? In: B. Calcott, R. Joyce, & K. Sterelny (eds.). *Signalling, Commitment & Cooperation*. MIT Press.

Heyes, C. (2016). Who Knows? Metacognitive Social Learning Strategies. *Trends in Cognitive Sciences*. 20 (3). pp. 1–10.

Heyes, C. & Pearce, J.M. (2015). Not-so-social learning strategies. *Proceedings of the Royal Society B: Biological Sciences*. 282 (1802). pp. 20141709.

Heyes, C.M. (1994). Social learning in animals: categories and mechanisms. *Biological reviews of the Cambridge Philosophical Society*. 69 (2). pp. 207–31.

Heyes, C.M. (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences*. 21. pp. 101–114.

Heyes, C.M. (2012b). What's social about social learning? *Journal of Comparative Psychology*. 126 (2). pp. 193–202.

Heyes, C.M. & Dawson, G.R. (1990). A demonstration of observational learning in rats using a bidirectional control. *The Quarterly Journal of Experimental Psychology. B, Comparative and Physiological Psychology*. 42 (1). pp. 59–71.

Heyes, C.M. & Ray, E.D. (2000). What Is the Significance of Imitation in Animals? *Advances in the Study of Behavior*. 29. pp. 215–245.

Hill, K.R., Walker, R.S., Bozicević, M., Eder, J., Headland, T., Hewlett, B., Hurtado, a M., Marlowe, F., Wiessner, P. & Wood, B. (2011). Co-residence patterns in hunter-gatherer societies show unique human social structure. *Science*. 331 (6022). pp. 1286–9.

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).

Hobaiter, C., Poisot, T., Zuberbühler, K., Hoppitt, W. & Gruber, T. (2014). Social network analysis shows direct evidence for social transmission of tool use in wild chimpanzees. *PLoS Biology*. 12 (9). pp. e1001960.

Hopper, L., Spiteri, a, Lambeth, S., Schapiro, S., Horner, V. & Whiten, a (2007). Experimental studies of traditions and underlying transmission processes in chimpanzees. *Animal Behaviour*. 73 (6). pp. 1021–1032.

Hopper, L.M. (2010). 'Ghost' experiments and the dissection of social learning in humans and animals. *Biological Reviews*. 85 (4). pp. 685–701.

Hopper, L.M., Flynn, E.G., Wood, L.A.N. & Whiten, A. (2010). Observational learning of tool use in children: Investigating cultural spread through diffusion chains and learning mechanisms through ghost displays. *Journal of Experimental Child Psychology*. 106 (1). pp. 82–97.

Hopper, L.M., Lambeth, S.P. & Schapiro, S.J. (2012). An Evaluation of the Efficacy of Video Displays for Use With Chimpanzees (Pan troglodytes). *American Journal of Primatology*. 74 (5). pp. 442–449.

Hopper, L.M., Lambeth, S.P., Schapiro, S.J. & Whiten, A. (2008). Observational learning in chimpanzees and children studied through 'ghost' conditions. *Proceedings of the Royal Society B: Biological Sciences*. 275 (1636). pp. 835–40.

Hoppitt, W., Brown, G., Kendal, R., Rendell, L., Thornton, A, Webster, M. & Laland, K. (2008). Lessons from animal teaching. *Trends in Ecology & Evolution*. 23 (9). pp. 486–493.

Hoppitt, W. & Laland, K. (2013). *Social Learning: An Introduction to Mechanisms, Methods, and Models*. Princeton: Princeton University Press.

Hoppitt, W. & Laland, K. (2008). *Social Processes Influencing Learning in Animals: A Review of the Evidence*.

Hoppitt, W., Samson, J., Laland, K.N. & Thornton, A. (2012). Identification of learning mechanisms in a wild Meerkat population. *PLoS ONE*. 7 (8). pp. e42044.

Horner, V. & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (Pan troglodytes) and children (Homo sapiens). *Animal Cognition*. 8 (3). pp. 164–81.

Horner, V., Whiten, A., Flynn, E. & de Waal, F.B.M. (2006). Faithful replication of foraging techniques along cultural transmission chains by chimpanzees and children. *Proceedings of the National Academy of Sciences of the United States of America*. 103 (37). pp. 13878–83.

Horowitz, A.C. (2003). Do humans ape? Or do apes human? Imitation and intention in humans (Homo sapiens) and other animals. *Journal of Comparative Psychology*. 117 (3). pp. 325–336.

Hothorn T, Bretz F, Westfall P (2013). multcomp: Simultaneous Inference in General Parametric Models. R package version 1.3-1, URL http://CRAN.R-project.org/package=multcomp.

Huber, L., Range, F., Voelkl, B., Szucsich, A., Virányi, Z. & Miklosi, A. (2009). The evolution of imitation: what do the capacities of non-human animals tell us about the mechanisms of imitation? *Philosophical Transactions of the Royal Society of London: Biological Sciences*. 364 (1528). pp. 2299–309.

Humle, T. & Snowdon, C.T. (2008). Socially biased learning in the acquisition of a complex foraging task in juvenile cottontop tamarins, Saguinus oedipus. *Animal Behaviour*. 75 (1). pp. 267–277.

Hunt, G.R. & Gray, R.D. (2003). Diversification and cumulative evolution in New Caledonian crow tool manufacture. *Proceedings of the Royal Society B:*

Biological Sciences. 270 (1517). pp. 867–74.

Hurford, J.R. (2007). *The Origins of Meaning: Language in the Light of Evolution*. Oxford: Oxford University Press.

Ives, A.R., Midford, P.E. & Garland, T. (2007). Within-species variation and measurement error in phylogenetic comparative methods. *Systematic biology*. 56 (2). pp. 252–70.

Jones, K.E., Bielby, J., Cardillo, M., Fritz, S.A., O'Dell, J., Orme, C.D.L., Safi, K., Sechrest, W., Boakes, E.H., Carbone, C., Connolly, C., Cutts, M.J., Foster, J.K., Grenyer, R., Habib, M., Plaster, C.A., Price, S.A., Rigby, E.A., Rist, J., Teacher, A., Bininda-Emonds, O.R.P., Gittleman, J.L., Mace, G.M. & Purvis, A. (2009). PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals W. K. Michener (ed.). *Ecology*. 90 (9). pp. 2648–2648.

Kandler, A. & Laland, K.N. (2013). Tradeoffs between the strength of conformity and number of conformists in variable environments. *Journal of theoretical biology*. 332. pp. 191–202.

Kavanagh, L.C., Suhler, C.L., Churchland, P.S. & Winkielman, P. (2011). When It's an error to mirror: the surprising reputational costs of mimicry. *Psychological Science*. 22 (10). pp. 1274–6.

Kawai, M. (1965). Newly-acquired pre-cultural behavior of the natural troop of Japanese monkeys on Koshima islet. *Primates*. 6 (1). pp. 1–30.

Kendal, J., Giraldeau, L.-A. & Laland, K. (2009a). The evolution of social learning rules: payoff-biased and frequency-dependent biased transmission. *Journal of Theoretical Biology*. 260 (2). pp. 210–9.

Kendal, R., Coolen, I., van Bergen, Y. & Laland, K. (2005). Tradeoffs in the adaptive use of social and asocial learning. *Advances in the Study of Behavior*. 35. p.pp. 333–379.

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). pp. 65–72.

Kendal, R.L., Kendal, J.R., Hoppitt, W. & Laland, K.N. (2009b). Identifying social learning in animal populations: a new 'option-bias' method. *PloS one*. 4 (8). pp. e6541.

Kenward, B. (2012). Over-imitating preschoolers believe unnecessary actions are normative and enforce their performance by a third party. *Journal of Experimental Child Psychology*. 112 (2). pp. 195–207.

Kenward, B., Karlsson, M. & Persson, J. (2011). Over-imitation is better explained by norm learning than by distorted causal learning. *Proceedings of the Royal Society*

B: Biological Sciences. 278 (1709). p.pp. 1239–46.

Keupp, S., Bancken, C., Schillmöller, J., Rakoczy, H. & Behne, T. (2016). Rational over-imitation: Preschoolers consider material costs and copy causally irrelevant actions selectively. *Cognition*. 147. pp. 85–92.

Keupp, S., Behne, T. & Rakoczy, H. (2013). Why do children overimitate? Normativity is crucial. *Journal of Experimental Child Psychology*. 116 (2). pp. 392–406.

Keupp, S., Behne, T., Zachow, J., Kasbohm, A. & Rakoczy, H. (2015). Over-imitation is not automatic: Context sensitivity in children's overimitation and action interpretation of causally irrelevant actions. *Journal of Experimental Child Psychology*. 130. pp. 163–175.

Klein, R.G. (2000). Archeology and the evolution of human behavior. *Evolutionary Anthropology: Issues, News, and Reviews*. 9 (1). pp. 17–36.

Kline, M.A. & Boyd, R. (2010). Population size predicts technological complexity in Oceania. *Proceedings of the Royal Society B: Biological Sciences*. 277 (1693). pp. 2559–64.

Koopmans, R. & Rebers, S. (2009). Collective action in culturally similar and dissimilar groups: an experiment on parochialism, conditional cooperation, and their linkages. *Evolution and Human Behavior*. 30 (3). pp. 201–211.

Lakin, J.L. & Chartrand, T.L. (2003). Using Nonconscious Behavioral Mimicry to Create Affiliation and Rapport. *Psychological Science*. 14 (4). pp. 334–339.

Lakin, J.L., Chartrand, T.L. & Arkin, R.M. (2008). I too am just like you: Nonconscious mimicry as an automatic behavioural response to social exclusion. *Psychological Science*. 19 (8). pp. 816–822.

Laland, K.N. (2004). Social learning strategies. *Learning & Behavior*. 32 (1). pp. 4–14.

Laland, K.N. & Hoppitt, W. (2003). Do animals have culture? *Evolutionary Anthropology: Issues, News, and Reviews*. 12 (3). pp. 150–159.

Laland, K.N. & Janik, V.M. (2006). The animal cultures debate. *Trends in Ecology & Evolution*. 21 (10). pp. 542–7.

Laland, K.N. & Kendal, J.R. (2003). What the models say about social learning. In: D. Fragaszy & S. Perry (eds.). *The Biology of Traditions*. Cambridge: Cambridge University Press, pp. 33–55.

Laland, K.N., Richerson, P.J. & Boyd, R. (1993). Animal social learning: Toward a new theoretical approach. *Perspectives in Ethology*. 10. pp. 249–277.

Lamba, S. (2014). Social learning in cooperative dilemmas. *Proceedings of the Royal Society B: Biological Sciences*. 281 (1787). pp 20140417

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). pp. 14426–14430.

Leadbeater, E. & Chittka, L. (2007). Social learning in insects--from miniature brains to consensus building. *Current Biology*. 17 (16). pp. 703–13.

van Leeuwen, E.J.C., Kendal, R.L., Tennie, C. & Haun, D.B.M. (2015). Conformity and its look-a-likes. *Animal Behaviour*. 110. pp. e1–e4.

Legare, C.H. & Nielsen, M. (2015). Imitation and Innovation: The Dual Engines of Cultural Learning. *Trends in Cognitive Sciences*. 19 (11). pp. 688–99.

Legare, C.H., Wen, N.J., Herrmann, P.A. & Whitehouse, H. (2015). Imitative flexibility and the development of cultural learning. *Cognition*. 142. pp. 351–361.

Leighton, J., Bird, G., Orsini, C. & Heyes, C. (2010). Social attitudes modulate automatic imitation. *Journal of Experimental Social Psychology*. 46 (0). pp. 905–910.

Levy, G.D., Taylor, M.G. & Gelman, S.A. (1995). Traditional and evaluative aspects of flexibility in gender roles, social conventions, moral rules, and physical laws. *Child Development*. 66 (2). pp. 515–31.

Lewis, H.M. & Laland, K.N. (2012). Transmission fidelity is the key to the build-up of cumulative culture. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. 367 (1599). pp. 2171–80.

Likowski, K.U., Mühlberger, A., Seibt, B., Pauli, P. & Weyers, P. (2008). Modulation of facial mimicry by attitudes. *Journal of Experimental Social Psychology*. 44 (4). pp. 1065–1072.

Lonsdorf, E. V. (2006). What is the role of mothers in the acquisition of termite-fishing behaviors in wild chimpanzees (Pan troglodytes schweinfurthii)? *Animal Cognition*. 9 (1). pp. 36–46.

Lumsden, C.J. & Wilson, E.O. (1981). *GENES, MIND, AND CULTURE: THE COEVOLUTIONARY PROCESS*. Cambridge, MA: Harvard University Press.

Lyons, D.E., Damrosch, D.H., Lin, J.K., Macris, D.M. & Keil, F.C. (2011). The scope and limits of overimitation in the transmission of artefact culture. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. 366 (1567). pp. 1158–67.

Lyons, D.E., Young, A.G. & Keil, F.C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences of the United States of America*. 104 (50). pp. 19751–6.

Mace, R. (2014). Human behavioral ecology and its evil twin. *Behavioral Ecology*. 25 (3). pp. 443–449.

Maddux, W.W., Mullen, E. & Galinsky, A.D. (2008). Chameleons bake bigger pies and take bigger pieces: Strategic behavioral mimicry facilitates negotiation outcomes. *Journal of Experimental Social Psychology*. 44 (2). pp. 461–468.

Manning, A. (1979). *An Introduction to Animal Behaviour*. 3rd Ed. London: Edward Arnold.

Mason, W.A., Jones, A. & Goldstone, R.L. (2008). Propagation of innovations in networked groups. *Journal of Experimental Psychology*. 137 (3). pp. 422–33.

Mcbrearty, S. & Brooks, A.S. (2000). The revolution that wasn't: a new interpretation of the origin of modern human behavior. *Journal of Human Evolution*. 39 (5). pp. 453–563.

McGuigan, N. (2013). The influence of model status on the tendency of young children to over-imitate. *Journal of Experimental Child Psychology*. 116 (4). pp. 962–9.

McGuigan, N., Makinson, J. & Whiten, A. (2011). From over-imitation to super-copying: adults imitate causally irrelevant aspects of tool use with higher fidelity than young children. *British Journal of Psychology*. 102 (1). pp. 1–18.

McGuigan, N. & Robertson, S. (2015). The influence of peers on the tendency of 3- and 4-year-old children to over-imitate. *Journal of Experimental Child Psychology*. 136. pp. 42–54.

McGuigan, N., Whiten, A., Flynn, E. & Horner, V. (2007). Imitation of causally opaque versus causally transparent tool use by 3- and 5-year-old children. *Cognitive Development*. 22 (3). pp. 353–364.

McIntosh, D.N. (2006). Spontaneous facial mimicry, liking and emotional contagion. *Polish Psychological Bulletin*. 37 (1). pp. 31–42.

Mellars, P. (2005). The impossible coincidence. A single-species model for the origins of modern human behavior in Europe. *Evolutionary Anthropology: Issues, News, and Reviews*. 14 (1). pp. 12–27.

Meltzoff, A.N. (2007). 'Like me': a foundation for social cognition. *Developmental Science*. 10 (1). pp. 126–34.

Mesoudi, A. (2008). An experimental simulation of the 'copy-successful-individuals' cultural learning strategy: adaptive landscapes, producer–scrounger dynamics, and informational access costs. *Evolution and Human Behavior*. 29 (5). pp. 350–363.

Mesoudi, A., Chang, L., Dall, S.R.X. & Thornton, A. (2016). The Evolution of Individual and Cultural Variation in Social Learning. *Trends in Ecology & Evolution*. 31 (3). pp. 215–225.

Mesoudi, A., Chang, L., Murray, K. & Lu, H.J. (2015). 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). pp. 20142209.

Mesoudi, A. & Whiten, A. (2008). The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. 363 (1509). pp. 3489–3501.

Mesoudi, A., Whiten, A. & Dunbar, R. (2006). A bias for social information in human cultural transmission. *British Journal of Psychology*. 97 (3). pp. 405–23.

Moraru, C.A., Gomez, J.-C. & McGuigan, N. (2016). Developmental changes in the influence of conventional and instrumental cues on over-imitation in 3- to 6-year-old children. *Journal of Experimental Child Psychology*. 145. pp. 34–47.

Morgan, T.J.H. & Laland, K.N. (2012). The biological bases of conformity. *Frontiers in Neuroscience*. 6 (June). pp. 87.

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). pp. 653–62.

Morgan, T.J.H., Uomini, N.T., Rendell, L.E., Chouinard-Thuly, L., Street, S.E., Lewis, H.M., Cross, C.P., Evans, C., Kearney, R., de la Torre, I., Whiten, a. & Laland, K.N. (2015a). Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nature Communications*. 6. pp. 6029.

Morgan, T.J.H., Laland, K.N. & Harris, P.L. (2015b). The development of adaptive conformity in young children: effects of uncertainty and consensus. *Developmental Science*. 18 (4). pp. 511–24.

Nagell, K., Olguin, R.S. & Tomasello, M. (1993). Processes of social learning in the tool use of chimpanzees (Pan troglodytes) and human children (Homo sapiens). *Journal of Comparative Psychology*. 107 (2). pp. 174–86.

Nielsen, M. (2008). The social motivation for social learning. *Behavioral and Brain Sciences*. 31 (01). pp. 33.

Nielsen, M. & Blank, C. (2011). Imitation in young children: when who gets copied is more important than what gets copied. *Developmental Psychology*. 47 (4). pp. 1050–3.

Nielsen, M., Mushin, I., Tomaselli, K. & Whiten, A. (2015). Where culture takes hold: 'overimitation' and its flexible deployment in Western, Aboriginal, and Bushmen children. *Child Development*. 85 (6). pp. 2169–84.

Nielsen, M. & Tomaselli, K. (2010). Overimitation in Kalahari Bushman children and the origins of human cultural cognition. *Psychological Science*. 21 (5). pp. 729–36.

Nishi, A., Shirado, H., Rand, D.G. & Christakis, N.A. (2015). Inequality and visibility of wealth in experimental social networks. *Nature*. 526 (7573). pp. 426–429.

Noad, M.J., Cato, D.H., Bryden, M.M., Jenner, M.N. & Jenner, K.C. (2000). Cultural revolution in whale songs. *Nature*. 408 (6812). pp. 537.

Nowak, M. A, Tarnita, C.E. & Antal, T. (2010). Evolutionary dynamics in structured populations. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*. 365 (1537). pp. 19–30.

Oostenbroek, J. & Over, H. (2015). Young children contrast their behavior to that of out-group members. *Journal of Experimental Child Psychology*. 139. pp. 234–241.

Ottoni, E.B. & Izar, P. (2008). Capuchin monkey tool use: Overview and implications. *Evolutionary Anthropology: Issues, News, and Reviews*. 17 (4). pp. 171–178.

Over, H. & Carpenter, M. (2012). Putting the social into social learning: explaining both selectivity and fidelity in children's copying behavior. *Journal of Comparative Psychology*. 126 (2). pp. 182–92.

Over, H. & Carpenter, M. (2013). The Social Side of Imitation. *Child Development Perspectives*. 7 (1). pp. 6–11.

Pagel, M. (2012). *Wired for culture. The natural history of human cooperation*. Allen Lane.

Pike, T.W. & Laland, K.N. (2010). Conformist learning in nine-spined sticklebacks' foraging decisions. *Biology Letters*. 6 (4). pp. 466–8.

Postma, E. & Charmantier, A. (2007). What 'animal models' can and cannot tell ornithologists about the genetics of wild populations. *Journal of Ornithology*. 148 (S2). pp. 633–642.

Powell, A., Shennan, S. & Thomas, M.G. (2009). Late Pleistocene demography and the appearance of modern human behavior. *Science*. 324 (5932). pp. 1298–301.

Pradhan, G.R., Tennie, C. & van Schaik, C.P. (2012). Social organization and the evolution of cumulative technology in apes and hominins. *Journal of Human Evolution*. 63 (1). pp. 180–90.

Price, E. & Caldwell, C.A. (2007). Artificially generated cultural variation between two groups of captive monkeys, Colobus guereza kikuyuensis. *Behavioural Processes*. 74 (1). pp. 13–20.

Price, E.E., Caldwell, C. A. & Whiten, A. (2010). Comparative cultural cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*. 1 (1). pp. 23–31.

Price, E.E., Lambeth, S.P., Schapiro, S.J. & Whiten, A. (2009). A potent effect of observational learning on chimpanzee tool construction. *Proceedings of the Royal Society B: Biological Sciences*. 276 (1671). pp. 3377–83.

Quinn, G. & Keough, M. (2007). *Experimental design and data analysis for biologists*. 6th Ed. Cambridge: Cambridge University Press.

Quinn, J.L., Charmantier, A., Garant, D. & Sheldon, B.C. (2006). Data depth, data completeness, and their influence on quantitative genetic estimation in two contrasting bird populations. *Journal of Evolutionary Biology*. 19 (3). pp. 994–1002.

Rand, D.G., Arbesman, S. & Christakis, N. A. (2011). Dynamic social networks promote cooperation in experiments with humans. *Proceedings of the National Academy of Sciences*. 108 (48). pp. 19193-19198

Rand, D.G., Dreber, A., Ellingsen, T., Fudenberg, D. & Nowak, M.A. (2009). positive interactions promote public cooperation. *Science*. 763. pp. 1272.

Rand, D.G., Greene, J.D. & Nowak, M.A. (2012). Spontaneous giving and calculated greed. *Nature*. 489 (7416). pp. 427–30.

Range, F., Huber, L. & Heyes, C. (2011). Automatic imitation in dogs. *Proceedings of the Royal Society B: Biological Sciences*. 278 (1703). pp. 211–217.

Range, F., Viranyi, Z. & Huber, L. (2007). Selective Imitation in Domestic Dogs. *Current Biology*. 17 (10). pp. 868–872.

Reader, S.M., Hager, Y. & Laland, K.N. (2011). The evolution of primate general and cultural intelligence. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*. 366 (1567). pp. 1017–27.

Reader, S.M. & Laland, K.N. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proceedings of the National Academy of Sciences of the United States of America*. 99 (7). pp. 4436–41.

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M.W., Fogarty, L., Ghirlanda, S., Lillicrap, T. & Laland, K.N. (2010a). Why copy others? Insights from the social learning strategies tournament. *Science*. 328 (5975). pp. 208–13.

Rendell, L., Boyd, R., Enquist, M., Feldman, M.W., Fogarty, L. & Laland, K.N. (2011a). How copying affects the amount, evenness and persistence of cultural knowledge: insights from the social learning strategies tournament. *Philosophical transactions of the Royal Society of London: Biological Sciences*. 366 (1567). pp. 1118–28.

Rendell, L., Fogarty, L., Hoppitt, W.J.E., Morgan, T.J.H., Webster, M.M. & Laland, K.N. (2011b). Cognitive culture: Theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences*. 15 (2). pp. 68–76.

Rendell, L., Fogarty, L. & Laland, K.N. (2010b). Rogers' paradox recast and resolved: population structure and the evolution of social learning strategies. *Evolution*. 64 (2). pp. 534–48.

Richerson, P., Baldini, R., Bell, A., Demps, K., Frost, K., Hillis, V., Mathew, S., Newton, E., Narr, N., Newson, L., Ross, C., Smaldino, P., Waring, T. & Zefferman, M. (2016). Cultural Group Selection Plays an Essential Role in

Explaining Human Cooperation: A Sketch of the Evidence. *The Behavioral and Brain Sciences*. In press. pp. 1–71.

Rigamonti, M.M., Custance, D.M., Previde, E.P. & Spiezio, C. (2005). Testing for Localized Stimulus Enhancement and Object Movement Reenactment in Pig-Tailed Macaques (Macaca nemestrina) and Young Children (Homo sapiens). *Journal of Comparative Psychology*. 119 (3). pp. 257–272.

Rogers, A. (1988). Does biology constrain culture. *American Anthropologist*. 90 (4). pp. 819–831.

Roitblat, H. (1998). Mechanisms of imitation: The relabeled story. *Behavioral and Brain Sciences*. 21. pp. 701–702.

Sachs, J.L., Mueller, U.G., Wilcox, T.P. & Bull, J.J. (2004). The evolution of cooperation. *The Quarterly Review of Biology*. 79 (2). pp. 135–60.

Saggerson, A.L., George, D.N. & Honey, R.C. (2005). Imitative learning of stimulus-response and response-outcome associations in pigeons. *Journal of Experimental Psychology*. 31 (3). pp. 289–300.

van Schaik, C.P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C.D., Singleton, I., Suzuki, A., Utami, S.S. & Merrill, M. (2003). Orangutan cultures and the evolution of material culture. *Science*. 299 (5603). pp. 102–5.

Scott-Phillips, T. (2014). *Speaking Our Minds: Why human communication is different, and how language evolved to make it special*. London: Palgrave Macmillan.

Seed, A. & Byrne, R. (2010). Animal Tool-Use. *Current Biology*. 20 (23). pp. 1032–1039.

Seed, A.M., Call, J., Emery, N.J. & Clayton, N.S. (2009). Chimpanzees solve the trap problem when the confound of tool-use is removed. *Journal of Experimental Psychology: Animal Behavior Processes*. 35 (1). pp. 23–34.

Seston, R. & Kelemen, D. (2014). Children's Conformity When Acquiring Novel Conventions: The Case of Artifacts. *Journal of Cognition and Development*. 15 (4). pp. 569–583.

Shea, N. (2009). Imitation as an inheritance system. *Philosophical transactions of the Royal Society of London: Biological sciences*. 364 (1528). pp. 2429–43.

Shettleworth, S. (1998). *Cognition, evolution, and behaviour*. New York: Oxford University Press.

Shumaker, R., Walkup, K. & Beck, B. (2011). *Animal tool behavior: the use and manufacture of tools by animals*. Baltimore, MD: Johns Hopkins University Press.

Slagsvold, T. & Wiebe, K.L. (2007). Learning the ecological niche. *Proceedings of the Royal Society B: Biological Sciences*. 274 (1606). pp. 19–23.

Spence, K. (1937). Experimental studies of learning and higher mental processes in infrahuman primates. *Psychological Bulletin*. 34. pp. 806–850.

Stel, L.L.E., Blascovich, J.I.M., Mccall, C., Mastop, J., Van Baaren, R. & Vonk, R. (2010). Mimicking disliked others : Effects of a priori liking on the mimicry-liking link. *European Journal of Social Psychology*. 40 (5). pp. 867–880.

Sterelny, K. (2003). *Thought in a Hostile World: The Evolution of Human Cognition*. Oxford: Blackwell.

Stoinski, T.S., Drayton, L. a & Price, E.E. (2011). Evidence of social learning in black-and-white ruffed lemurs (Varecia variegata). *Biology Letters*. 7 (3). pp. 376–9.

Stoinski, T.S. & Whiten, A. (2003). Social learning by orangutans (Pongo abelii and Pongo pygmaeus) in a simulated food-processing task. *Journal of Comparative Psychology*. 117 (3). pp. 272–82.

Stoinski, T.S., Wrate, J.L., Ure, N. & Whiten, A. (2001). Imitative learning by captive western lowland gorillas (Gorilla gorilla gorilla) in a simulated food-processing task. *Journal of Comparative Psychology*. 115 (3). pp. 272–81.

Stout, D. (2002). Skill and Cognition in Stone Tool Production: An Ethnographic Case Study from Irian Jaya. *Current Anthropology*. 43 (5). pp. 693–722.

Stout, D. (2011). Stone toolmaking and the evolution of human culture and cognition. *Philosophical transactions of the Royal Society of London: Biological Sciences*. 366 (1567). pp. 1050–9.

Stout, D., Toth, N., Schick, K. & Chaminade, T. (2008). Neural correlates of Early Stone Age toolmaking: technology, language and cognition in human evolution. *Philosophical transactions of the Royal Society of London: Biological Sciences*. 363 (1499). pp. 1939–49.

Stringer, C. (2007). The Origin and Dispersal of Homo sapiens: Our Current State of Knowledge. In: P. Mellars, K. Boyle, C. Stringer, & O. Bar-Yosef (eds.). *Rethinking the Human Revolution: New Behavioural and Biological Perspectives on the Origin and Dispersal of Modern Humans*. McDonald Institute for Archaeological Research (December 30, 2007), pp. 15–21.

Van Swol, L. & Drury-Grogan, M. (submitted). The effects of shared opinions on nonverbal mimicry. *Sage Open Science*.

Van Swol, L.M. (2003). The Effects of Nonverbal Mirroring on Perceived Persuasiveness, Agreement with an Imitator, and Reciprocity in a Group Discussion. *Communication Research*. 30 (4). pp. 461–480.

Tanner, R., Ferraro, R., Chartrand, T., Bettman, J. & Van Baaren, R. (2008). Of chameleons and consumption: The impact of mimicry on choice and preferences. *Journal of Consumer Research*. 34 (April). pp. 754–766.

Tennie, C., Call, J. & Tomasello, M. (2006). Push or Pull: Imitation vs. Emulation in

Great Apes and Human Children. *Ethology*. 112 (12). pp. 1159–1169.

Tennie, C., Call, J. & Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of cumulative culture. *Philosophical transactions of the Royal Society of London: Biological Sciences*. 364 (1528). pp. 2405–15.

Thompson, B., Kirby, S. & Smith, K. (2016). Culture shapes the evolution of cognition. *Proceedings of the National Academy of Sciences*. *published early online*.

Thornton, A. & Malapert, A. (2009). The rise and fall of an arbitrary tradition: an experiment with wild meerkats. *Proceedings of the Royal Society B: Biological Sciences*. 276 (1660). pp. 1269–76.

Thorpe, W.H. (1963). *Learning and instinct in animals*. 2nd Ed. London: Methuen.

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). pp. 326–333.

Toelch, U., Bruce, M.J., Newson, L., Richerson, P.J. & Reader, S.M. (2014). Individual consistency and flexibility in human social information use. *Proceedings of the Royal Society B: Biological Sciences*. 218 (20132864). pp. 2–9.

Tomasello, M. (1990). Cultural transmission in the tool use and communicatory signaling of chimpanzees? In: S. T. Parker & K. R. Gibson (eds.). *'Language' and intelligence in monkeys and apes: Comparative Developmental Perspectives*. Cambridge: Cambridge University Press, pp. 274–311.

Tomasello, M. (1996). Do apes ape? In: C. M. Heyes & B. Galef (eds.). *Social learning in animals: The roots of culture*. New York: Academic Press, pp. 319–346.

Tomasello, M. (1998). Emulation learning and cultural learning. *Behavioral and Brain Sciences*. 21 (5). pp. 703–704.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M. & Call, J. (1997). *Primate Cognition*. New York: Oxford University Press.

Tomasello, M. & Herrmann, E. (2010). Ape and Human Cognition: What's the Difference? *Current Directions in Psychological Science*. 19 (1). pp. 3–8.

Tomasello, M., Kruger, A.C. & Ratner, H.H. (1993). Cultural Learning. *Behavioral and Brain Sciences*. 16 (3). pp. 495–511.

Torsten Hothorn, Frank Bretz and Peter Westfall (2008). Simultaneous Inference in General Parametric Models. Biometrical Journal 50(3), 346-363.

Tuttle, R.H. (2001). On culture and traditional chimpanzees. *Current Anthropology*. 42 (3). pp. 407–408.

Vaesen, K., Collard, M., Cosgrove, R. & Roebroeks, W. (2016). Population size does not explain past changes in cultural complexity. *Proceedings of the National Academy of Sciences*. *early online*.

de Villemereuil, P., Gimenez, O. & Doligez, B. (2013). Comparing parent-offspring regression with frequentist and Bayesian animal models to estimate heritability in wild populations: a simulation study for Gaussian and binary traits R. *Methods in Ecology and Evolution*. 4 (3). pp. 260–275.

Visalberghi, E. & Addessi, E. (2000). Seeing group members eating a familiar food enhances the acceptance of novel foods in capuchin monkeys. *Animal Behaviour*. 60 (1). pp. 69–76.

Voelkl, B. & Huber, L. (2007). Imitation as faithful copying of a novel technique in marmoset monkeys. *PloS ONE*. 2 (7). pp. e611.

Voelkl, B. & Huber, L. (2000). True imitation in marmosets. *Animal Behaviour*. 60 (2). pp. 195–202.

van de Waal, E., Borgeaud, C. & Whiten, A. (2013). Potent social learning and conformity shape a wild primate's foraging decisions. *Science*. 340 (6131). pp. 483–5.

de Waal, F., Boesch, C., Horner, V. & Whiten, A. (2008). Comparing social skills of children and apes. *Science*. 319 (5863). pp. 569.

Want, S.C. & Harris, P.L. (2001). Learning from other people's mistakes: causal understanding in learning to use a tool. *Child Development*. 72 (2). pp. 431–43.

Warneken, F., Chen, F. & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*. 77 (3). pp. 640–63.

Warner, R.R. (1988). Traditionality of mating-site preferences in a coral reef fish. *Nature*. 335 (6192). pp. 719–721.

Wasielewski, H. (2014). Imitation Is Necessary for Cumulative Cultural Evolution in an Unfamiliar, Opaque Task. *Human Nature*. 25 (1). pp. 161–179.

Watson-Jones, R.E., Whitehouse, H. & Legare, C.H. (2015). In-Group Ostracism Increases High-Fidelity Imitation in Early Childhood. *Psychological Science*. 27 (1). pp 34-42

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). pp. 2869–76.

West, S.A., Griffin, A.S. & Gardner, A. (2007). Social semantics: Altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of*

Evolutionary Biology. 20 (2). pp. 415–432.

West, S.A., El Mouden, C. & Gardner, A. (2011). Sixteen common misconceptions about the evolution of cooperation in humans. *Evolution and Human Behavior*. 32 (4) pp. 231–262.

Whiten, A., Goodall, J., McGrew, W.C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C.E., Wrangham, R.W. & Boesch, C. (1999). Cultures in chimpanzees. *Nature*. 399 (6737). pp. 682–5.

Whiten, A. (1998). Imitation of the sequential structure of actions by chimpanzees (Pan troglodytes). *Journal of comparative psychology*. 112 (3). pp. 270–81.

Whiten, A., Custance, D.M., Gomez, J.C., Teixidor, P. & Bard, K.A. (1996). Imitative learning of artificial fruit processing in children (Homo sapiens) and chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*. 110 (1). pp. 3–14.

Whiten, A., Flynn, E., Brown, K. & Lee, T. (2006). Imitation of hierarchical action structure by young children. *Developmental Science*. 9 (6). pp. 574–582.

Whiten, A. & Ham, R. (1992). On the Nature and Evolution of Imitation in the Animal Kingdom: Reappraisal of a Century of Research. *Advances in the Study of Behavior*. 21. pp. 239–283.

Whiten, A., Horner, V., Litchfield, C. A & Marshall-Pescini, S. (2004). How do apes ape? *Learning & Behavior*. 32 (1). pp. 36–52.

Whiten, A., Horner, V. & de Waal, F.B.M. (2005). Conformity to cultural norms of tool use in chimpanzees. *Nature*. 437 (7059). p.pp. 737–40.

Whiten, A., McGuigan, N., Marshall-Pescini, S. & Hopper, L.M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical transactions of the Royal Society of London: Biological Sciences*. 364 (1528). pp. 2417–28.

Whiten, A. & van Schaik, C.P. (2007). The evolution of animal 'cultures' and social intelligence. *Philosophical transactions of the Royal Society of London: Biological Sciences*. 362 (1480). pp. 603–20.

Wilks, M., Collier-Baker, E. & Nielsen, M. (2015). Preschool children favor copying a successful individual over an unsuccessful group. *Developmental Science*. 18(6). pp1014-2

Williamson, R.A., Meltzoff, A.N. & Markman, E.M. (2008). Prior experiences and perceived efficacy influence 3-year-olds' imitation. *Developmental Psychology*. 44 (1). pp. 275–85.

Wilson, A.C. (1985). The Molecular Basis of Evolution. *Scientific American*. 253 (4). pp. 164–173.

Wisdom, T.N., Song, X. & Goldstone, R.L. (2013). Social learning strategies in

networked groups. *Cognitive Science*. 37 (8). pp. 1383–425.

Wittenberger, J. (1981). *Animal Social Behaviour*. Boston: Duxbury Press.

Wolf, M., Kurvers, R.H.J.M., Ward, A.J.W., Krause, S. & Krause, J. (2013). Accurate decisions in an uncertain world: collective cognition increases true positives while decreasing false positives. *Proceedings of the Royal Society B: Biological Sciences*. 280 (1756). pp. 2777.

Wood, L.A., Kendal, R.L. & Flynn, E.G. (2012). Context-dependent model-based biases in cultural transmission: children's imitation is affected by model age over model knowledge state. *Evolution and Human Behavior*. 33 (4). 387-394.

Wood, L.A., Kendal, R.L. & Flynn, E.G. (2013). Copy me or copy you? The effect of prior experience on social learning. *Cognition*. 127 (2). pp. 203–13.

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. pp. 190–202.

Yabar, Y., Johnston, L., Miles, L. & Peace, V. (2006). Implicit Behavioral Mimicry: Investigating the Impact of Group Membership. *Journal of Nonverbal Behavior*. 30 (3). pp. 97–113.

Yamamoto, S., Humle, T. & Tanaka, M. (2013). Basis for Cumulative Cultural Evolution in Chimpanzees: Social Learning of a More Efficient Tool-Use Technique. *PLoS ONE*. 8 (1). pp. e55768

Zeileis, A., Kleiber, C., Jackman, S. (2008). Regression Models for Count Data in R. Journal of Statistical Software 27(8).

Zentall, T.R. (2001). Imitation in animals: Evidence, function, and mechanisms. *Cybernetics and Systems*. 32 (1-2). pp. 53–96.

Zwirner, E. & Thornton, A. (2015). Cognitive requirements of cumulative culture: teaching is useful but not essential. *Scientific Reports*. 5 (16781).