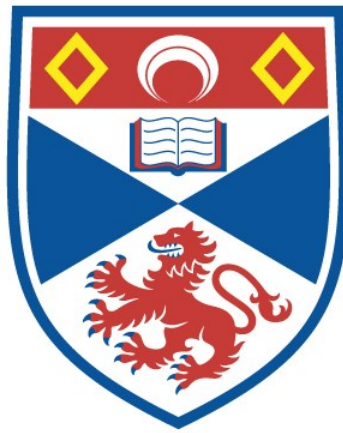


# **EXPERIMENTAL STUDIES OF HUMAN SOCIAL LEARNING AND ITS EVOLUTION**

**Thomas J.H. Morgan**

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



**2014**

**Full metadata for this item is available in  
St Andrews Research Repository  
at:**

**<http://research-repository.st-andrews.ac.uk/>**

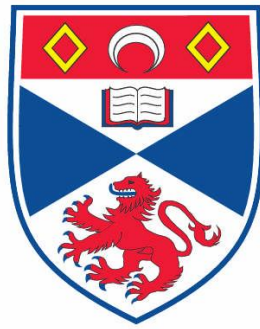
**Please use this identifier to cite or link to this item:**

**<http://hdl.handle.net/10023/4443>**

**This item is protected by original copyright**

**This item is licensed under a  
Creative Commons Licence**

# Experimental Studies of Human Social Learning and its Evolution



Thomas J. H. Morgan

This thesis is submitted in partial fulfilment to the  
University of St. Andrews for the degree of  
**Doctor of Philosophy**

October, 2013

### **Candidate's declarations:**

I, Thomas Morgan, hereby certify that this thesis, which is approximately 50,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September, 2009 and as a candidate for the degree of Doctor of Philosophy in September, 2009; the higher study for which this is a record was carried out in the University of St Andrews between 2009 and 2013.

Date: 7/10/2013, 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 Doctor of Philosophy in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date: 7/10/2013, signature of supervisor:

### **Permission for electronic 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 electronic publication of this thesis:

Embargo on both all of printed copy and electronic copy for the same fixed period of 2 years on the grounds that publication would be commercially damaging to the researcher, or to the supervisor, or the University.

Date 7/10/2013, signature of candidate

signature of supervisor

## **Contributions:**

### **Chapters 1 and 2:**

Some of this work is available as a published manuscript:

Morgan, T. J. H., and Laland, K. N. (2012). The Biological Bases of Conformity. *Front. Neurosci.* 6, 1–7.

### **Chapter 3:**

This work was carried out with Prof. Kevin Laland, Dr. Luke Rendell, Dr. William Hoppitt and Dr. Micke Ehn. LR wrote the code for experiment 3.1, ME provided assistance with the coding of the remainder of the experiments and WH provided assistance with the analysis of the data. Experimental sessions were run by myself, LR and ME. All other work was carried out by myself. This work is available as a published manuscript:

Morgan, T., Rendell, L., Ehn, M., Hoppitt, W., and Laland, K. (2011). The evolutionary basis of human social learning. *Proceedings of the Royal Society B: Biological Sciences*.

### **Chapter 4:**

This work was carried out with Prof. Paul Harris and Prof. Kevin Laland. PH provided advice concerning experimental design, both KL and PH offered input on interpretation and preparation of a manuscript. All other work was carried out by myself.

### **Chapter 5:**

This work was carried out with Dr. Natalie Uomini, Dr. Luke Rendell, Laura Thuly, Sally Street, Dr. Hannah Lewis, Dr. Catherine Cross, Cara Evans, Ronan Kearney, Dr. Ignacio de la Torre, Prof. Andrew Whiten and Prof. Kevin Laland. NU, LR, IdIT and KL provided input on experimental design, NU and LR assisted with carrying-out two pilot studies, NU, LR, LT, SS, HL, CC and CE assisted with carrying-out the experiment, IdIT and NU triple coded flakes, RK coded the videos and all authors offered input on interpretation and preparation of a manuscript. All other work was carried out by myself.

My work was supported by an ERC Advanced Grant to Kevin Laland (EVOCULTURE, Ref: 232823).

## Acknowledgements

Foremost I would like to thank my supervisor Kevin Laland for the mix of encouragement and scrutiny that separated the scientific wheat from my mental chaff. I would also like to thank my co-supervisor Luke Rendell for advice and assistance on numerous fronts, occasional devils advocacy and a disturbing interest in raw data.

I have also received valuable council from several other individuals and in particular I would like to thank Natalie Uomini for advice and assistance throughout the stone knapping project. In addition I would like to thank Paul Harris for assisting throughout the child project and for allowing me the opportunity to spend time with his group. I would also like to thank Andrew Whiten and Ignacio de la Torre for their input into the stone knapping project.

I would also like to thank Katherine Meacham for the innumerable forms of help she offered during my time in the Laland Lab, both with experiments, but also in minimising the pain and confusion caused by having to fill out forms.

Several members of the Laland lab have helped me tremendously during my time in St Andrews. I would like to thank Kate Cross for informal supervision, help with experiments, tea and chocolate; Will Hoppitt for his patience as he watched me wrap my head around BUGS; Micke Ehn for the same with java; Sally Street, Cara Evans, Hannah Lewis and Laura Chouinard-Thuly for helping run experiments; and all above along with Alice Cowie, Riva Riley, Camille Troisi, Michal Arbilly, Sarah Benson-Amram, Neeltje Boogert, Glenna Evans, Stuart Murray, Ana Navarette, Mike Webster, Nicola Atton, Laurel Fogarty, Lewis Dean and Theoni Photopoulou for being part of a wonderful working environment. Although only informally part of the Laland Lab, I would like to thank Ronan Kearney for giving up too much of his time to watch videos of apes bashing rocks together.

I would also like to thank the Media Lab staff of the Bute Building and the Discovery Centre staff at the Boston Museum of Science for giving me a space in which to carry out experiments; Darko Odic, Justin Halberda and Lisa Feigensen for advice on how to create numerical discrimination tasks and John & Val Lord for driving 2 tonnes of flint to Scotland and teaching us to knap.

Finally I would like to thank Elisabeth Morgan for moral support and sanity maintenance throughout.

## Summary

Human culture is unique in its scope and complexity and is underpinned by the social transmission of information. Successful individuals will use both social and asocial information effectively. Evolutionary theory suggests that social learning should be guided by evolved learning rules that dictate when individuals rely on social information, a literature which I review across **Chapters 1 and 2**, with the emphasis of chapter 2 being on conformist transmission. In this thesis I present experimental investigations of the existence and adaptive value of several such strategies in both adults (**Chapter 3**) and young children (**Chapter 4**). In all cases I find strong evidence for the existence of such biases and show that they act to increase the accuracy of decisions. In particular I show individuals are highly sensitive to even small majorities within a group of demonstrators. The youngest children (age 3) however, show little sensitivity to social information and do not use it effectively. In **Chapter 5** I present an investigation into the role of social learning in the evolution of hominin lithic technology. I conclude that even the earliest hominin flaking technology is poorly transmitted through observation alone and so the widespread and longstanding persistence of such tools implies some form of teaching. Furthermore, I conclude that the stable transmission of more complex technologies would likely require teaching, and potentially symbolic communication. I also postulate a co-evolution of stone tools and complex communication and teaching. In **Chapter 6** I conclude that the cultural evolutionary approach, focussing on the evolutionary consequences of social information use and treating culture as a system of inheritance partially independent of genes, seems successful in increasing our understanding of the evolution of social learning.

## Contents

<b>1 - Introduction</b>	<b>0</b>
1.1 - An introduction to strategic social learning	1
1.2 - Social psychology & conformity	2
1.3 - Developmental psychology and trust	5
1.4 - Cultural evolution and social learning	6
1.4.1 - Pure social and asocial learning	7
1.4.2 - Critical and conditional social learning	9
1.4.3 - Uncertainty	10
1.4.4 - Payoff-, prestige- and reliability-biased social learning	12
1.4.5 - Dominance-, age-, size- and kin-biased social learning	16
1.4.6 - Unbiased transmission	20
1.5 - Neurological studies	21
1.6 - Summary	24
<b>2 - An Overview of Conformist Transmission</b>	<b>25</b>
2.1 - Introduction to conformist transmission	26
2.2 - Cultural evolution modelling	28
2.2.1 - Anti-conformist and unbiased transmission	30
2.3 - Cultural evolution experiments	34
2.3.1 - Social psychological studies	36
2.4 - Animal experimentation	37
2.5 - The development of conformist transmission	39
2.6 - The neurobiology of conformist transmission	40
2.7 - Conclusions	41
<b>3 - An Empirical Study of Social Learning Strategies in Adults</b>	<b>43</b>
3.1 - Introduction	45
3.2 - Methods	47
3.2.1 - General methods	47
3.2.1.1 - Participants and apparatus	48
3.2.1.2 - Procedure	49
3.2.1.3 - General design	49
3.2.1.4 - Methods of information presentation	50
3.2.2 - The tasks	51
3.2.2.1 - The foraging task	51
3.2.2.2 - The mental rotation task	52
3.2.2.3 - The length estimation task	53
3.2.2.4 - Audio tasks	53
3.2.3 - Types of social information	55
3.2.4 - The experiments	56
3.2.4.1 - Experiment 3.1	56
3.2.4.2 - Experiment 3.2	57
3.2.4.3 - Experiment 3.3	57
3.2.4.4 - Experiment 3.4	58
3.3 - Analyses	59
3.3.1 - Experiment 3.1	60
3.3.2 - Experiment 3.2	61
3.3.3 - Experiment 3.3	62

3.3.4 - Experiment 3.4	63
3.4 - Results	64
3.4.1 - Experiment 3.1	64
3.4.2 - Experiment 3.2	64
3.4.3 - Experiment 3.3	70
3.4.4 - Experiment 3.4	71
3.5 - Discussion	72
3.5.1 - The number of demonstrators	72
3.5.2 - Participant confidence	73
3.5.3 - The cost of asocial information	74
3.5.4 - Changes in social learning across trials	74
3.5.5 - Different tasks	75
3.5.6 - Payoff-biased copying	76
3.5.7 - Frequency dependent copying & conformist transmission	78
3.5.8 - A bias against social information	79
3.5.9 - Conclusions	81
<b>4 - An Empirical Study of Social Learning Strategies in Young Children</b>	<b>82</b>
4.1 - Introduction	84
4.1.1 - Copy-when-uncertain	84
4.1.2 - Conformist transmission	85
4.1.3 - Outstanding questions	86
4.2 - Methods	88
4.2.1 - General methods	88
4.2.2 - Subjects & apparatus	89
4.2.3 - The task	89
4.2.4 - The social information	91
4.2.5 - Procedure	91
4.3 - Analysis	92
4.3.1 - Priors	95
4.4 - Results	96
4.5 - Discussion	101
4.5.1 - Consensus and conformist transmission	101
4.5.2 - Confidence and information quality	103
4.5.3 - The adaptive value of children's social learning	105
4.5.4 - Asocial performance	106
4.5.5 - Conclusions	107
<b>5 - An Empirical Case-Study of the Social Transmission of Oldowan Lithic Flaking Technology</b>	<b>109</b>
5.1 - Introduction	111
5.2 - Methods	113
5.2.1 - General methods	113
5.2.2 - Apparatus & set-up	114
5.2.3 - Procedure	115
5.2.4 - Learning conditions	117
5.2.5 - Participant payment	119
5.2.6 - Recorded variables	121
5.3 - Coding	121
5.3.1 - Individual flakes	121
5.3.2 - Videos	126
5.3.3 - Language	127



5.4 - Analyses	131
5.4.1 - General analyses	131
5.4.2 - Individual flakes	131
5.4.3 - Aggregate flakes	132
5.4.4 - Flake identification	133
5.4.5 - Core remaining	133
5.4.6 - Videos	134
5.4.7 - Language	135
5.5 - Results	136
5.5.1 - General results	136
5.5.2 - Individual flakes	137
5.5.3 - Aggregate flakes	141
5.5.4 - Flake identification	149
5.5.5 - Core remaining	152
5.5.6 - Videos	154
5.5.7 - Language	162
5.6 - Discussion	177
5.6.1 - Aggregate flakes	177
5.6.2 - Core remaining	179
5.6.3 - Videos	180
5.6.4 - Language	181
5.6.5 - Flake identification	182
5.6.6 - Individual flakes	183
5.6.7 - Effects of position along chains	184
5.6.8 - Effects of core mass	186
5.6.9 - General Discussion	187
5.6.9.1 - The importance of teaching	187
5.6.9.2 - The co-evolution of stone tools and communication	190
5.6.9.3 - Implications for Oldowan stasis and the appearance of Acheulean technology	192
5.6.10 - Extensions and improvements	193
5.6.11 - Conclusions	194
<b>6 - General Discussion</b>	<b>196</b>
6.1 - Introduction	197
6.2 - Summary of findings	197
6.3 - Broader considerations	200
6.3.1 - The status of adaptive learning rules	200
6.3.2 - Conformist transmission	201
6.3.3 - Gene-culture co-evolution	203
6.4 - Further work	204
6.4.1 - The evolved basis of learning rules	204
6.4.2 - The neural basis of learning rules	206
6.4.3 - Normative motivations for copying	206
6.4.4 - A quantitative mechanism for social learning	207
6.5 - Conclusion	208

## Chapter 1

# **Introduction**

## **1.1 – An introduction to strategic social learning**

All animals must make decisions throughout their lives, many of which, such as foraging or mating decisions, will have fitness consequences. Additionally, all animals will encounter other individuals both of the same species and of different species. For group living animals, including humans, the presence of conspecifics is a constant feature of their environment. As all individuals within a population likely face many of the same decisions, then the choices they make, and the information they have access to, will potentially be of value to individuals other than themselves. Furthermore, the information of others may be highly attractive to observers as it offers a potentially very cheap way to gain knowledge of the environment (Boyd and Richerson, 1985). For example, when deciding where to forage, individuals can investigate foraging sites themselves. However, this exploration comes with a cost, in terms of the time and energy spent investigating several options, and also an increased risk of predation, as the potential foraging sites may not be safe locations (Boyd and Richerson, 1985). If an individual could access information already gathered by other individuals then they would be able to learn about potential foraging sites indirectly, without having to pay the cost of directly gathering this information. However, social information comes with its own problems as an observer cannot be sure of its quality and it may be out of date, or inappropriate (Boyd and Richerson, 1985; Kameda and Nakanishi, 2002). Given this, an optimal decision maker will successfully combine information they collect independently of others (asocial information) with information available to them through the behaviour or products of other individuals (social information). This raises the possibility that individuals may be guided by behavioural rules that dictate how asocial and social sources of information are

combined, referred to as "social learning strategies" (Laland, 2004; Rendell et al., 2011) or "transmission biases" (Boyd and Richerson, 1985; Henrich and McElreath, 2003). The investigation of such rules need not assume they are genetically specified, indeed, it is perfectly possible that they are themselves products of learning. However, due to the fitness consequences of such rules, it is likely that natural selection would have fashioned their operation, regardless of the extent to which such rules are genetically underpinned (Laland, 2004; Morgan et al., 2011). In the sections that follow I first summarise three fields of relevant investigations into the use of social and asocial information, namely, social psychology, developmental psychology and cultural evolution. For the rest of this chapter I describe in more detail many of the studied learning rules, highlighting multi-disciplinary support and outstanding questions.

## **1.2 - Social psychology & conformity**

Some of the earliest scientific investigations into use of social information were carried out by social psychologists during the 20<sup>th</sup> century, and were focused very much on the social contexts that would cause individuals to forgo their own opinions in favour of those of others (Asch, 1956; Sherif, 1935; Jenness, 1932). In an extremely influential paper, Solomon Asch (1956) described the observation that adults would willingly abandon their own perceptual judgement in a very simple visual task when faced with a group of confederates who disagreed with them. In this study adult participants were presented with three lines of different lengths and instructed to identify the largest of the lines. The lines were clearly different in length making identification of the longest line easy. Also present were a number of

confederates who appeared to the participant to be taking part in the study as they themselves were. When the group was asked to respond, the confederates gave incorrect decisions such that by the time the genuine participant made a decision they had received social information conflicting with their own perceptual information. Asch termed the adoption of the group decision at the expense of perceptual information “conformity”, supposing the deference to the group norm to be driven by a desire to receive social approval. Such a finding has been replicated a huge number of times across age groups and cultures and a large number of factors that influence whether or not individuals conform have been identified including group size (Asch, 1956; Bond, 2005), task difficulty and importance (Baron et al., 1996), culture (Bond and Smith, 1996) motivation (Griskevicius et al., 2006) and mood (Tong et al., 2008). Whilst social psychology, replete with empirical data, has successfully identified many factors influencing when individuals adopt the decisions of others, it has struggled to unify such findings into a single theoretical framework. Perhaps the most successful attempt is Social Impact Theory (Latane, 1981; Latane and Wolf, 1981; Nowak et al., 1990), which characterizes social influence as a force, analogous to a physical force such as electro-magnetism, that acts on an individual. Factors proposed to influence the magnitude of this force are its strength (determined by factors such as the age and status of the source), immediacy (proximity in space-time of the source to the observer), and the number of people in the group to which the observer is exposed. Social Impact Theory can effectively explain the diminishing effect of increasing the number of confederates in the Asch experiments (Latane, 1981), and was also extended to cases where a majority conflicted with a minority (Latane and Wolf, 1981). However, its variables of strength and immediacy - precisely that which distinguished it from other models (e.g., Tanford and Penrod, 1984) - came up against

conflicting empirical findings and where effects were found they were typically of a very low magnitude (Mullen, 1986; Jackson, 1986; Mullen, 1985). Furthermore, social psychology does not typically operate within an evolutionary framework. One ramification of this is that these theories were largely based on studies involving the adoption of arbitrary or bizarre group decisions. As a result the ability of these studies to explain social influence more generally, particularly in the context of evolution, is limited. Accordingly, the ambitions of theories of social influence from social psychology, although valuable contributions to the study of social learning, were never fully realised.

Nonetheless, social influence theories were very successful in accounting for group size effects. Moreover, social psychology is also the source of a valuable distinction between "informational" and "normative" motivations for conforming to a group norm (Deutsch and Gerard, 1955). This distinction came about as researchers attempted to understand why their subjects were conforming to clearly incorrect decisions. They argued for two goals on the part of the subject, one to be correct, but a second to earn positive appraisal from others through agreement. The former is an informational goal, the latter a normative goal. As the simplicity of the task in the Asch experiments seems to preclude an informational goal, it has been argued that the subjects were conforming in order to achieve a normative reward, received by being in agreement with group mates. Surprisingly, Deutsch and Gerard (1955) found that some subjects would still choose the clearly incorrect answer even when they made their decision in the absence of confederates. They took this to mean that the confederates were also exerting some informational influence and that the subjects may really have believed

the group decisions. An alternative explanation is that, even when apparently isolated, individuals may find normative tendencies hard to resist.

### **1.3 - Developmental psychology and trust**

There have also been many studies of social learning in children, referred to as "trust", from developmental psychology (Harris, 2007, 2012), which have attempted to understand under what conditions children will rely on the testimony of other individuals. Like social psychology, developmental psychology does not operate within an evolutionary framework and so does not typically consider the adaptive value of such behaviour. Arguably, this is potentially less of a problem for studies of young children because the predictions of evolutionary models may not apply to the behaviour of children which may still be developing. However, this is contentious and many behavioural adaptations specific to childhood are thought to exist. For example, it has been argued that children show a special sensitivity to adult cues that signify an intent to communicate relevant knowledge such that children learn effectively (Gergely et al., 2007).

Either way, researchers can still make predictions about the behaviour of young children, and there are at least three reasons to anticipate that developmental changes in social learning may occur. Firstly, cognitive and behavioural strategies that are adaptive for adults may not be for children, in which cases adaptive age-specific social learning strategies may have been favoured by selection. For example, children tend to interact with smaller groups (e.g. family, carers) than adults, in which case some strategies, such as conformist transmission (see **Chapter 2**), might be less

effective . Secondly, if the development of social learning strategies is contingent on experience of social interactions, we should expect strategy usage to change as a child's social horizon expands. This is because, being part of a larger social group will lead to a greater variety of social interactions. Thirdly, very young children may be incapable of some forms of computation necessary to implement adaptive social learning strategies (e.g. computing the majority behavior which is a requisite for conformist transmission, see **Chapter 2**), rendering their decision making suboptimal prior to passing a threshold developmental stage. The three cases described above offer different (although not mutually exclusive) explanations for changes in behavior across development; in the first case alternative strategies have been selected, in the second they emerge depending on relevant experience, whilst in the third case they are explained by the presence/absence of other cognitive processes.

#### **1.4 - Cultural evolution and social learning**

Starting in the 1970s, biologists interested in the possibility of evolved learning rules, and operating within an evolutionary framework, began to study not only the existence, but also the adaptive value, of social learning under the name of “transmission biases” and “social learning strategies” (Boyd and Richerson, 1985; Laland, 2004; Cavalli-Sforza and Feldman, 1981; Henrich and McElreath, 2003). The nascent fields of cultural evolution and gene-culture co-evolution distinguished themselves from other evolutionary approaches to understanding human behaviour and culture in that they treated culture as an evolving system at least partially independent of genes. This stance allowed the evolution of maladaptive cultural traits and the relationship between genes and culture was reciprocal in that culture could

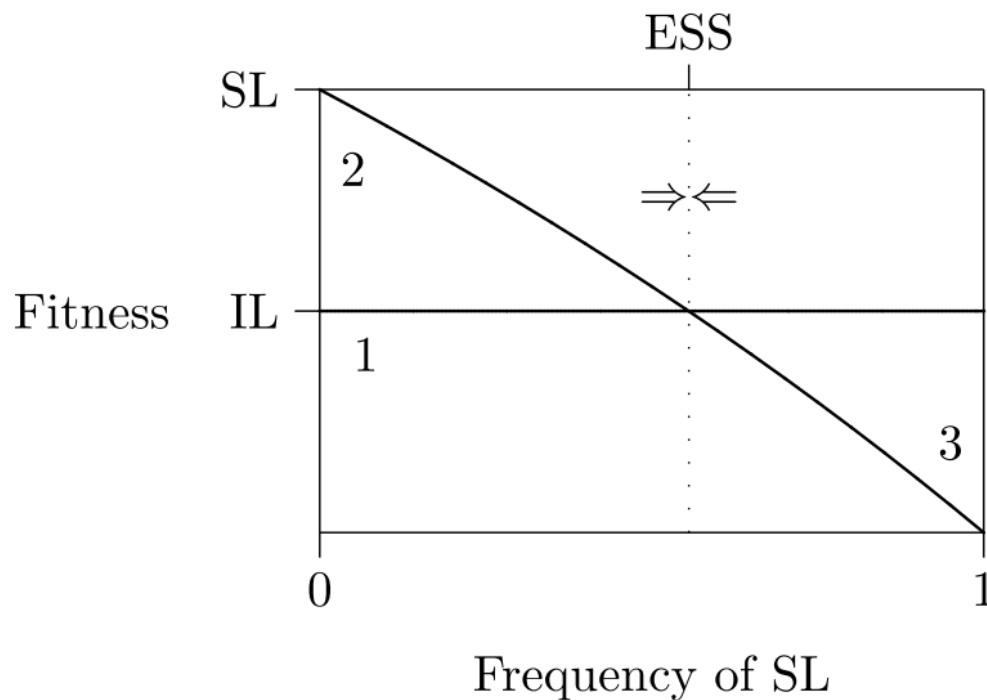


impact upon genetic evolution. In this case, social transmission becomes the means of heritability for the cultural system, the equivalent of genetic inheritance for genetic evolution. For much of the history of cultural evolution and gene-culture co-evolution a theoretical approach was adopted, with mathematical models being developed to investigate the evolution and persistence of transmission biases. However more recently empirical studies have begun to test the predictions of these models. For the rest of this chapter I will now describe the various forms of social learning biases that have been considered and illustrate the support for these derived from theoretical models and experiments with adults, children and non-human animals. One particular learning rule – conformist transmission – is of sufficient complexity and importance to my thesis that rather than deal with it here I dedicate the entire second chapter to its consideration.

#### **1.4.1 – Pure social and asocial learning**

Studies of social learning were heavily influenced by the theoretical finding now commonly referred to as Rogers' paradox. Anthropologist, Alan Rogers (1988) modelled the evolution of a mixed population of pure asocial learners and pure social learners - individuals who were entirely reliant on asocial or social information respectively - inhabiting a changing environment, in order to investigate when it paid to rely on each form of knowledge gain. The social learners were "unbiased" in their transmission in that they copied another member of the population chosen at random. His key finding was that social learning, despite being assumed to be cheaper than asocial learning, did not raise the average population fitness. This is because as social learning grows in prevalence the amount of up to date information entering the

population through asocial learning diminishes and so the behaviour of social learners becomes increasingly out of sync with the environment leading to poor decision making. Over time the population evolves to an equilibrium at which all individuals have the same fitness, and where the average fitness of all individuals is equivalent to that of a population of pure asocial learners (see **Fig. 1.1**). This result, although not strictly paradoxical, was viewed by many researchers as surprising because social learning at the time was typically considered to be adaptive and believed to be behind human population growth (Richerson and Boyd, 2005). Furthermore, the equilibrium frequency reached in the model was far lower than the ubiquitous presence of social learning observed in human society implied. This finding spurred other theoretical work to examine social learning rules that might increase the fitness of social learners above that of asocial learners.



**Figure 1.1**

Figure taken from Rogers (1988). As the frequency of social learning (SL) increases it's fitness decreases. Meanwhile, the fitness of asocial, or individual, learning (IL) does not depend on the frequency of social learning. At the ESS the fitness of the two strategies is equal, and the same as a population consisting only of asocial learners.

#### 1.4.2 – Critical and conditional social learning

Two examples of such strategies are "critical" and "conditional" social learning (Enquist et al., 2007). Critical social learners first learn socially, taking advantage of the relatively cheap cost of social information. However, unlike Rogers' pure social learners they then evaluate the quality of the information they received and if it is lacking they go on to gather their own information asocially. Such a strategy was found to evolve over a wide range of conditions, including regions of parameter space typically considered representative of real environments where social learning is cheap and asocial learning more costly, but reasonably effective. Conditional social

learning, is the inverse of critical social learning, where asocial learning is the default and social learning only used if the former fails. Enquist et al. (2007) found that this evolves in opposite regions of parameter space to critical social learning – where asocial learning is cheaper than social learning, although these areas are perhaps less typical of natural environments. Although these strategies do lead to an increase in overall population fitness beyond that of pure asocial learners, an outcome that was not possible in Rogers' model, a difficulty comes with the means by which individuals are able to evaluate the quality of any information they receive. Such investigation might be regarded as itself a form of asocial learning, and so any critical social learner who sticks with the decision of the social information cannot be considered a pure social learner. However, more problematic is whether such evaluation is at all possible in the real world. For many decisions, for example deciding which type of crop to plant, it is not possible to evaluate the quality of the decision without having fully committed to it (i.e., planting a crop and seeing how it goes), by which point it is no longer possible to gather some more information and reverse your decision.

### **1.4.3 - Uncertainty**

Conditional social learning bears a great similarity to the social learning strategy “copy-when-uncertain” by which individuals are more likely to adopt the decisions of others when they feel uncertain in their own abilities or knowledge (Laland, 2004). The existence of a bias to copy when uncertain has been a key assumption of theoretical models of cultural evolution. Boyd and Richerson (1988) modelled individuals in an environment that changed between two possible states. They

postulated than when individual experience was unsatisfactory (i.e. it left them uncertain) individuals should adopt the decisions of others. Although such uncertainty may be generated through an attempt at collecting asocial information that results in insufficient evidence to make a decision, it is also possible that it could be due to inference from poor performance on the same task on previous occasions or on non-identical but related tasks. Thus, an individual who makes a poor mating decision one year may come to doubt their ability to identify high-quality mates and so be more inclined to copy the decisions of others when required to make another decision on the next year. Evidence consistent with the deployment of a copy-when-uncertain strategy can be seen in a study of children presented with a numerical discrimination task (Odic et al., 2012). In this study half of the children were given the questions in increasing order of difficulty (i.e., easy questions first), whilst the other half were given the questions in decreasing order of difficulty. The children given the harder questions first expressed lower confidence in their abilities, even on the easy questions, compared to the children given the easier questions first. However, the crucial final step, to show that young children use their confidence to guide reliance on social information was not taken in their study. Elsewhere, a bias to copy when uncertain has received empirical support across a variety of non-human taxa including rats (Galef et al., 2008), gerbils (Forkman, 1991), capuchin monkeys (Visalberghi and Frigaszy, 1995), ants (Grüter et al., 2011) and nine-spined sticklebacks (Kendal et al., 2005). For example, van Bergen et al. (2004) presented nine-spined sticklebacks with two patches at which they could choose to feed, one of which was rich and the other poor. Across seven training days which patch was rich varied such that some groups of fish received unreliable information where one patch was rich only slightly more often than the other, whilst other groups received more reliable information where one

patch was rich more consistently. After this, trained fish were shown a group of fish foraging at the two patches, but where the patch that had most frequently been rich during training was now poor. Finally, the observer fish were required to make a decision again. van Bergen et al. (2004) found that fish given reliable asocial information showed a strong preference towards the patch that was typically reliable during training. However, when the prior information had been unreliable this preference was greatly reduced and the fish were influenced to a greater extent by the social information. Given this, it seems plausible that confidence guided social learning is likely to be a part of human social learning, indeed, this finding is reported in **Chapter 3**.

#### **1.4.4 – Payoff-, prestige- and reliability-biased social learning**

Payoff biased social learning refers to any form of selective social learning where an individual's social learning is guided by the payoffs to themselves or to other individuals (Schlag, 1998; Kendal et al., 2009). Game-theoretic analyses have indicated that strategies where an individual's use of social information is guided by their own pay-off ("proportional reservation"), the pay-off to demonstrators ("proportional observation"), or the difference between the two ("proportional imitation") can all be highly effective (Schlag, 1998, 1999). Indeed, there is good evidence that humans and other animals are sensitive to such information and do use it to direct social learning (Mesoudi and O'Brien, 2008; Caldwell and Millen, 2008a; Mesoudi, 2008; Pike et al., 2010; Apesteguia et al., 2007). For example, in a computer-based arrowhead design task, in which participants could alter four different arrowhead parameters before receiving feedback on the efficacy of their design,

participants were found to selectively copy the design of the individual who was performing the best in the group (Mesoudi and O'Brien, 2008). However, although the value of payoff-biased copying seems clear in a general sense, which form it takes in humans or other species remains unclear with current data insufficient to resolve between the possibilities mentioned above. More generally, payoff-biased social learning suffers from a similar problem to the quality investigation required by critical social learning, in that precise information of payoffs may not be available to individuals until well after decisions are made, if at all. Given this, although they are certainly successful strategies in models and experiments where such information is available, and while there will certainly be some circumstances where payoffs are immediately visible, it remains unclear how often such strategies are a possibility in real populations.

“Prestige” is defined as noncoerced, interindividual, within-group human status asymmetry and prestige-biased social learning has been subject to attention by cultural evolutionists (Henrich and Gil-White, 2001), whereby individuals exhibit a general tendency to copy the decisions of an individual who had been successful in the relevant domain and who is afforded prestige because of this. Prestige can be distinguished from dominance through the stipulation that prestige is noncoercive, whereas dominance is coercive. Such copying of "expert" or generally successful individuals may be easier to implement than relying on knowledge of the payoffs associated with particular decisions and so could be a widespread phenomenon. For example, in a Fijian population a handful of *yalewa vuku*, or wise-women, had a greatly disproportionate impact on the cultural evolution of the population (Henrich and Henrich, 2010). Although there are some claims of prestige-biased copying in

non-human species (see **section 1.4.5**), such claims do not meet Henrich & Gil-White's definition. Prestige develops across multiple interactions that may be significantly separated in time and space, and requires individuals to monitor each other's long term success, which may be beyond the cognitive capacities of most non-human animals.

Although prestige- and payoff-based learning are not often discussed in studies of childhood social learning, the child's assessment of the reliability of sources of social information is. Reliability fits very well with the notion of payoffs and prestige as it is based on the repeated observations of the performance of a potential informant. For example, Koenig et al. (Koenig et al., 2004), carried out a study of three and four-year-old children in which they were exposed to a "reliable" and an "unreliable" informant by showing two individuals correctly and incorrectly naming familiar objects. When subsequently presented with a novel object and asked to name it the children were more likely to use the word suggested by the previously reliable informant than the name suggested by the previously unreliable informant. Furthermore, this selective trust is spontaneous, in that children do not need to be prompted to evaluate reliability, and it is sensitive to the frequencies of correct and incorrect answers, in that children can discern between two individuals both of whom make some correct and some incorrect judgements based on the frequency by which they make correct judgements (Pasquini et al., 2007). Corriveau & Harris (2009) carried out an experiment that required children to choose between reliable and unreliable informants sometime after watching them make decisions and found that assessments of reliability persist and guide social learning for at least a week after the initial reliability assessments. Despite this, the selective social learning of young



children does have limitations. For example, Koenig & Harris (2005) found that, when shown reliable and unreliable informants, 3-year-olds, unlike 4-year-olds, are unable to predict that an unreliable informant will get future decisions wrong and do not seek out information from a reliable informant, although children of both ages were more likely to adopt the suggestions of the reliable informant than those of the unreliable informant. The appearance of these behaviours in 4-year-olds strongly suggests the development of a deeper understanding of reliability at this age. However, Wood et al. (2012) investigated the social transmission of the solution to a simple extractive foraging task with 5-year-olds. They found that the age of the informant was a better predictor of copying than the informant's declared state of knowledge concerning the task, with children favouring adult informants over children. This could reflect increased reliance on simple heuristics (i.e., copy older individuals) at the expense of more nuanced, but cognitively demanding, strategies that involve knowledge state estimations. However, it could also reflect the typical roles of adults and children that 5-year-olds will be familiar with - i.e., adults are knowledgeable and are the sources of social information. It is also possible that the two responses could be task specific - perhaps on motor tasks adults are perceived as more reliable informants and so exert greater influence. Another example of social learning changing across development can be seen when children interact with a reliable informant and an ignorant informant – i.e., an informant who expresses no knowledge on a matter, as opposed to an unreliable informant who makes incorrect decisions (Koenig and Harris, 2005). In this case, when the previously ignorant informant does offer an opinion, 4-year-olds, but not 3-year-olds, favour the previously reliable informant over the ignorant informant. Children are also capable of responding to indirect cues of reliability, for example, children show a persistent

distrust of individuals who seem unreliable on the basis that they were a lone dissenter against a group of other informants who were in agreement (Corriveau et al., 2009a), and children are also more likely to adopt someone's opinions when they are endorsed by other individuals (Fusaro and Harris, 2008).

In addition to these studies with humans there is also evidence for payoff-guided social learning from other species. For example, nine-spined sticklebacks have been observed to copy the foraging patch decisions of more successful fish (Pike et al., 2010). This study even had the necessary precision to distinguish between Schlag's (1998) three learning rules and found that stickleback behaviour matched "proportional observation". Given the above, payoff-guided learning, at least in some form, may be a widespread phenomenon.

#### **1.4.5 - Dominance-, age-, size- and kin-biased social learning**

Whilst prestige seems rare outside of humans, dominance systems are taxonomically widespread and can also guide social learning. For example, the claim of prestige biased copying in chimpanzees (Horner et al., 2010) based on an experiment where individuals chose to copy more "prestigious" demonstrators over less "prestigious" demonstrators both of whom were exhibiting different solutions to a foraging problem, is more parsimoniously considered an example of dominance-based copying as the individuals cited as the most "prestigious" were the most dominant. Although dominance-biased social learning has not been investigated with evolutionary models, it could be an adaptive strategy if individuals achieved their social dominance through access to high-quality information. However, in many social species dominance is

achieved through alliances (e.g., Maestriperi, 2007; Engh et al., 2000) in which case dominant individual may provide good information concerning who to form alliances with, although not necessarily concerning other decisions. Evidence from both birds and monkeys suggests dominance-biased copying may be widespread as an indirect result of the behaviour of low-ranking individuals being inhibited by others (Drea and Wallen, 1999; Nicol and Pope, 1994). Thus, if low ranking individuals are unable to provide social information then it will come disproportionately from higher ranking individuals.

In addition to becoming socially dominant, successful individuals are likely to grow to a larger size than unsuccessful individuals due to the increased access to resources that their success affords them. Thus, again, we might predict a bias to copy larger individuals might guide observers towards adaptive outcomes. Data for such a bias can be seen in the guppy (Duffy et al., 2009), where fish were more likely to adopt the foraging decisions of larger demonstrator individuals than smaller demonstrator individuals (although in this study size and age were confounded).

Another bias that could guide social learning is age. Such a bias is likely to be adaptive to the extent that older individuals have had more time to acquire valuable information, but also have demonstrated, through their survival, that they have not been overly reliant on poor information. Thus, one might surmise that any remaining old individuals are more likely than young individuals to be in the possession of high-quality information. This is backed up by experimental evidence from guppies (Amlacher and Dugatkin, 2005) where female mating preferences were influenced to a greater degree by the equivalent decisions of older conspecific females than they

were by the decisions of younger conspecific females (see also Duffy et al., 2009). There is also evidence for such a bias in humans, for example, in the case of the Fijian population, although *yalewa vuku* or wise-women were particularly influential, to a lesser extent, so were *qase* - non-specific elders (Henrich and Henrich, 2010). Other studies have found that young birds (Biondi et al., 2010) and chimpanzees (Biro et al., 2003, 2006) are more reliant on social information than older conspecifics.

Individuals may also show a bias to copy the decisions of kin. There are at least three reasons as to why this is likely to be adaptive and hence favoured by selection. Firstly, kin members are more likely than non kin to have spent time in the same environment as the individual. Thus, the information they possess may be of greater relevance than that offered by unrelated individuals, including being suitable to their genotype. Secondly, kin are readily available, accessible and tolerant to close proximity. Third, due to the accrual of indirect fitness, individuals may be more likely to donate information to their kin, or to make such information more readily available (i.e., to teach them, Fogarty et al., 2011). Although reliable data on kinship in non-human animals is hard to acquire, there is a lot of circumstantial data, particularly in primates, which supports the existence of kin directed social learning (Reader, 2000). In a human, Fijian population, in addition to the influence of *yalewa vuku* and *qase* (wise-women and elders), food taboos were identified as primarily learned from mothers, grandmothers or mothers-in-law (Henrich and Henrich, 2010). In the case of vertical transmission from parents to offspring, however, it is possible that such a copying bias may not be the result of a psychological mechanism, but instead the result of behavioural practices concerning childrearing. In species with parental care of offspring, including humans, it is typically the parents with whom offspring

interact most frequently. Thus, even unbiased social learning might be expected to lead to a greater cultural transmission between parent-offspring pairs than between other members of the population.

The reliance of children on information provided by their mothers has been investigated by developmental psychologists. Children who were identified at 15 months as having a secure, ambivalent or avoidant relationship with their mother were found at age 5 to rely differently on information provided by their mother (Corriveau et al., 2009b). Avoidant infants did not discriminate between the information offered by their mother or a stranger, whilst secure infants showed a preference for their mother's opinion and ambivalent infants an even greater preference. A preference for information from familiar individuals has also been observed, children between the ages of 3 and 5 have also been shown to preferentially adopt the decisions of a familiar caregiver over an unfamiliar caregiver (Corriveau and Harris, 2009a). When subsequently observing the familiar caregiver make a number of errors the 4- and 5-year-olds, but not the 3-year-olds, became more reliant on the unfamiliar caregiver (Corriveau and Harris, 2009a), suggesting that trust of familiar individuals develops earlier than the ability to estimate, and guide trust according to, reliability. Although the trust of familiar caregivers is unlikely to have an adaptive value in terms of indirect fitness, it is still the case that a familiar individual is likely to have encountered similar environments to the observer and so is likely to provide more accurate local information. In support of this suggestion, children have also been shown to prefer information provided by a speaker with a native accent over information provided by a speaker with a non-native accent (Kinzler et al., 2011). Presumably this is because a native accent is providing a cue to the children that the

speaker, coming from the local area, is likely well informed about local conditions. Mirroring this, Chen et al. (2013) found that 4- to 7-year-old children prefer the information provided by an informant of the same race as themselves. However, the authors found that, children would ignore the advice of an informant of the same race as themselves if that individual was the lone dissenter against a consensus of a different racial composition.

#### **1.4.6 - Unbiased transmission**

Despite all the evidence for strategic social learning described above, there is evidence that under some conditions copying of randomly chosen individuals can be favoured (Rendell et al., 2010). Furthermore, the population level distributions of many human cultural traits, such as baby names, dog breeds, or the decorative motifs on pottery from Neolithic German farming settlements, obey a power law, similar to that expected under random copying (Bentley et al., 2004; Hahn and Bentley, 2003; Bentley et al., 2007; Shennan and Wilkinson, 2001). This does not necessarily imply that individuals are (or were) making these decisions randomly. Even if individuals were making decisions with great thought, as is likely the case when naming children, if different individuals use different criteria to make a decision, the net effect on the population distribution will be similar to that seen under random copying. However, more powerful methods to identify group level patterns in decision making affecting variant frequency have been developed (Kandler and Shennan, 2013), and when applied to the case of Neolithic German pottery they find evidence that variant frequency was being affected by such collective forces (Kandler and Shennan, 2013).

Future work with these methods will help identify the extent to which unbiased transmission is useful to understand cultural traits at the population level.

### **1.5 - Neurological studies**

Whilst the aforementioned studies have made ground in isolating the contexts that elicit social learning, this is just one aspect of the immediate cause of this behaviour. A complete understanding requires some knowledge of what goes on in the brains of individuals as they use social information. However, the neural-level processes underlying social learning are comparatively poorly understood. Nonetheless, recent studies have come up with several relevant findings. Firstly, studies using both mental rotation tasks (Berns et al., 2005) and auditory tasks (Berns et al., 2010) have found that social information affected neural activity in the relatively low level processing areas associated with each task, in addition to areas distinct from these perceptual decision making circuits, suggesting that the social information was affecting the perception of the subjects as well as their decision making, a possibility raised by Asch in the interpretation of his findings (Asch, 1956). In addition to this, ventral striatum activity in a music choice task (Campbell-Meiklejohn et al., 2010) suggests that the social information was directly affecting the perceived value of different songs. These findings are consistent with the idea that social and asocial sources of information are integrated at early stages of processing, however, the low temporal resolution of fMRI limits the strength of such a conclusion. Finally, Mason et al. (Mason et al., 2009) exposed subjects to symbols that received positive, negative or no social labelling. Exposure to a socially marked symbol resulted in activity in the medial prefrontal cortex, irrespective of whether it was positively or negatively

marked, whilst activity in the caudate coded the valence of the social marking. These findings suggest that it is through the integration of activity in these two areas that individuals distinguish between positively and negatively socially marked stimuli.

Neurobiological experiments have also shown that the traditional dichotomy between normative and informational motivations to use social information may not accurately represent why individuals use social information. For example, experiments typically attempt to explain subject behaviour in terms of one or the other source of influence and even posit different behavioural responses when subjects are influenced by one or the other (Campbell and Fairey, 1989). However, evidence from neurological studies provide evidence that the two processes may be unavoidably intertwined. For example, Berns et al. (2005) found increased brain activity when subjects disagreed with human participants as opposed to computers, despite the fact that the task was not obviously normative in nature. However, this could be a result of subjects placing more weight on human responses than those of computers. Other studies have found activity in areas strongly suggestive of a normative response, including the amygdala, an area associated with emotional load, suggesting that subjects found their being in disagreement with others stressful (Klucharev et al., 2009). However Edelson et al. (2011) carried out a study in which participants were asked to recall details of a video they had watched several weeks before having been more recently given false information. They found that the amygdala showed increased activity only when the false information purportedly came from other individuals (as opposed to computers) *and* when the subject subsequently altered their long-term responses accordingly. That no such activity was seen when behavioural adjustments were temporary suggests the activity was related to memory modification and so an emotional load may not have



been involved. However, the activity was only seen when the information came from humans, indicative of a normative aspect. Berns et al. (2010) found similar activity in the insula, an area associated with anxiety and ostracism, whilst Campbell-Meiklejohn et al. (2010) found activity in the lateral prefrontal cortex, an area linked with reputation management, this activity also predicted subsequent behavioural adjustments to the group norm. The similarities between the response to human and computer decisions could be interpreted as subjects anthropomorphising the computers, or alternatively treating human demonstrators as machines. Such findings imply that if researchers are to understand social information use, including conformity and conformist transmission, at the behavioural level it may be insufficient to consider it in light of either informational or normative influence in isolation as they may not be distinct processes at the neural level. A more complete theory of social decision making may need to include both informational and normative influences, with variable payoffs associated with getting the correct answer and fitting in with group mates. Experiments where both are included and their relative strengths manipulated could help our understanding of how the two interact. From this perspective, social decision making typically involves a maximisation of reward, taking into account the information provided by others on the group norm and the level consensus behind it, the expected cost of deviating from such a consensus, the individual's own information concerning the task, the information provided by others concerning the task and the expected cost of making an incorrect decision. To proceed with our understanding of social decision making it may be necessary to combine the above elements into a single theoretical framework and to stop thinking of behaviour in terms of either normative and social influences.

## **1.6 – Summary**

As can be seen from the above, the multiple approaches of social and developmental psychology along with cultural evolution have been highly instructive in increasing our understanding of human and non-human animal social learning strategies. This is both in terms of what factors influence when an individual will use social information, but also, in the case of cultural evolution, in terms of setting the behaviour within an evolutionary context and highlighting its taxonomic distribution and adaptive potential. Thus the cultural evolutionary approach has the promise of succeeding where social psychology struggled: in the attempt to unify social learning into a single conceptual framework. However, despite its successes, the approach is still in its infancy, with empirical studies somewhat sparse and several outstanding questions remaining. In particular, neurobiological data is largely absent. In the next chapter I shall turn my focus to one particular social learning strategy, conformist transmission, in great detail. This strategy merits a lengthy treatment as it will be central to much of the data I subsequently present, but also because research into conformist transmission has resulted in many conflicting findings – both empirical and theoretical.

## Chapter 2

# **An Overview of Conformist Transmission**

## 2.1 - Introduction to conformist transmission

In 1963, ethologist Niko Tinbergen argued that to fully understand behaviour in biology it is necessary to consider it from four different perspectives, that of its history, ontogeny, function and causation. Whilst heuristics such as these have the potential to constrain research as much as assist it (Laland et al., 2011) the realisation of answers to Tinbergen's four questions is often a very helpful target for research, encouraging a broad perspective on behaviour and fostering interdisciplinary approaches. The focus of this chapter will be a review of work concerning conformist transmission, a topic that has received considerable attention from at least three of these perspectives and is central to my thesis. Conformist transmission is a learning rule by which individuals are disproportionately likely to adopt the decisions of majorities, at the expense of minorities (Boyd and Richerson, 1985, p.206, see **Fig 2.1**). Consider the case of a naïve individual choosing between options A and B who is presented with 7 informants advocating option A and 3 informants advocating option B. In this case, the majority amongst the informants makes up 70% of the group. If the naïve individual were conformist they would have a greater than 70% chance of choosing option A (see **Fig. 2.1**). Although this learning rule is sometimes referred to as "conformity", to avoid confusion with work from social psychology where the term "conformity" has a different meaning, I shall refer to it as "conformist transmission", and shall refer to individuals who exhibit conformist transmission as "conformists". Across the rest of this chapter I summarise insights into conformist transmission derived from cultural evolution modelling, experimentation across a range of taxa and neurological studies. As the approaches of these fields do not map cleanly onto Tinbergen's questions, the chapter is organised

on a broadly disciplinary basis, but draws attention to which of Tinbergen's questions is being addressed in each case.

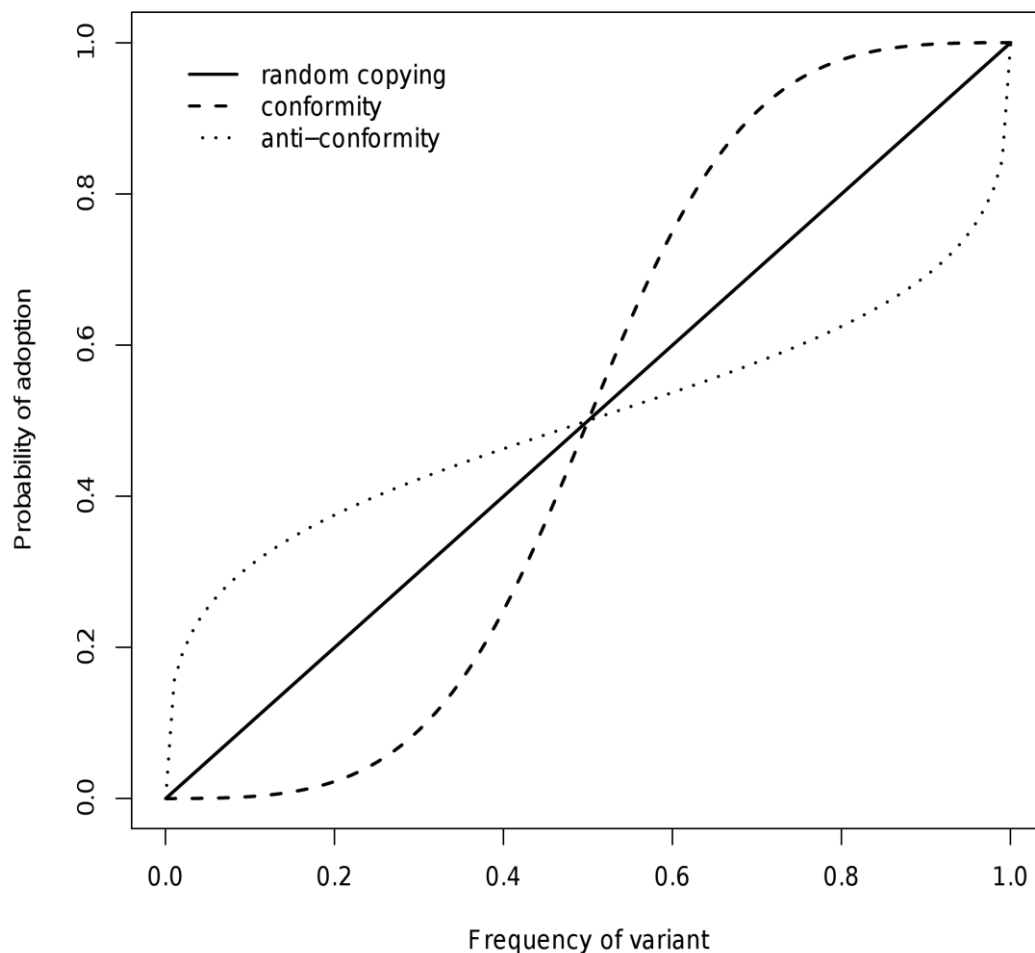


Figure 2.1

Conformist transmission (the dashed line) is one of several learning rules that result in the increasingly likely adoption of a trait as it increases in frequency. However, it differs from other such learning rules as the tendency towards the most popular trait is disproportionate given its frequency. A proportionate tendency, also known as "unbiased copying" and equivalent to random copying (solid line), results in a probability of adoption equal to trait frequency. Anti-conformist transmission (dotted line) resists the most popular choice and has the opposite population consequences to conformity. Of these frequency dependent rules, only conformist transmission generates homogenous group behaviour.

## 2.2 - Cultural evolution modelling

Mathematical models, studying the adaptive value (function *sensu* Tinbergen) of conformist transmission, have established a range of conditions under which conformist transmission is an effective strategy. For example, a model of naive migrant individuals in a spatially variable environment found that upon joining a new group in an unfamiliar environment, conformist transmission was a very effective means by which the migrant individual could accurately learn about their new environment because it helps individuals to home in on the locally adaptive behaviour (Boyd and Richerson, 1985). More recent work (Nakahashi et al., 2012) added to this, finding that spatial variation, errors in learning, and the number of options between which individuals choose, all favour the evolution of conformist transmission. This is because conformist transmission uses the decisions of a large group of individuals to identify potentially weak signals across multiple decisions. Both errors in learning and a larger number of options to choose between make each individual's decision less reliable, but when offered a group of individuals the correct option will still most likely be the most prevalent decision in the population and so conformist transmission will be a successful strategy. As with social learning more generally (Rogers, 1988), selection on conformist transmission is highly frequency dependent, with individual learners required to sustain conformists. Kandler & Laland (2013) found that very strong conformist transmission (where the tendency to copy small majorities is extremely strong) can only be sustained at low frequencies in a population, requiring a large number of asocial learners present to be successful. However, a weaker form of conformist transmission, where the bias towards the majority is only mildly disproportionate, requires fewer asocial learners to be successful and so can exist

successfully at much greater frequencies within a population. This highlights a feature of conformist transmission in that it can be an obstacle to the spread of innovations when rare (Eriksson et al., 2007). As new ideas must initially start at very low frequencies the prevalence of conformist transmission can act to block their spread. This is particularly problematic in temporally variable environments where the discovery and spread of new behaviours is essential to success (Eriksson et al., 2007; Nakahashi et al., 2012; Kandler and Laland, 2013). As weak conformist transmission hinders the spread of innovations less than strong conformist transmission, Kandler & Laland (2013) argue that conformist transmission is likely to be weak. Thus the extent to which conformity is expected to be adaptive is contested, but the evidence from theoretical models on balance leads us to expect a broad range of conditions under which it will be utilised.

Conformist transmission, as investigated in these models, fits well with an informational notion of conformity from social psychology (see **Chapter 1, section 1.2**) – people conform because it leads them to acquire valuable fitness-enhancing information. Nonetheless, conformist transmission could be associated with normative motivations for conforming too, as in this case the goal is to agree with as much of the population as possible and conformist transmission is an effective means by which to identify and adopt the modal opinion of the group (Boyd and Richerson, 1985; Richerson and Boyd, 2005). A second outcome of conformist transmission is the combination of stable intergroup heterogeneity and intragroup homogeneity, which fits well with patterns in human populations (Richerson and Boyd, 2005). When a population is divided into several groups such that individuals typically interact with their own group members, conformist transmission will lead each group to conform to

its own modal behaviour, which may well vary across groups due to chance or because of differences in environmental conditions (Boyd and Richerson, 1985). The end result of this process will be the emergence of cultural differences between groups of individuals. Once established, such group differences potentially allow a process of cultural group selection to occur (Boyd and Richerson, 1985; Richerson and Boyd, 2005; Henrich, 2004a). By this process, stable cultural differences between groups that lead to differential group success and cause groups possessing the most favourable cultural traits to out compete other groups that lack these traits, lead to the overall population becoming dominated by groups exhibiting favourable cultural traits. This process is more plausible than genetic group selection which is hampered by invasion of successful groups by individuals possessing genes that benefit the individual, but not the group. This is because, with cultural group selection, migrating individuals can adopt the culture of their new group thereby preserving group differences (Boyd and Richerson, 1985; Henrich, 2004a). As a result, cultural group selection has been argued to be behind the high levels of cooperation exhibited by human populations (Boyd and Richerson, 1985; Richerson and Boyd, 2005; Henrich, 2004a).

### **2.2.1 - Anti-conformist and unbiased transmission**

Conformist transmission is not the only frequency-dependent learning rule that has been considered. Anti-conformist transmission can be thought of as the opposite of conformist transmission as it is the disproportionately large influence of minorities on an individual (Cavalli-Sforza and Feldman, 1981). In the hypothetical scenario of 7



out of 10 demonstrators supporting option A, an anti-conformist individual would have a less than 70% chance of choosing option A (see **Fig. 2.2**).

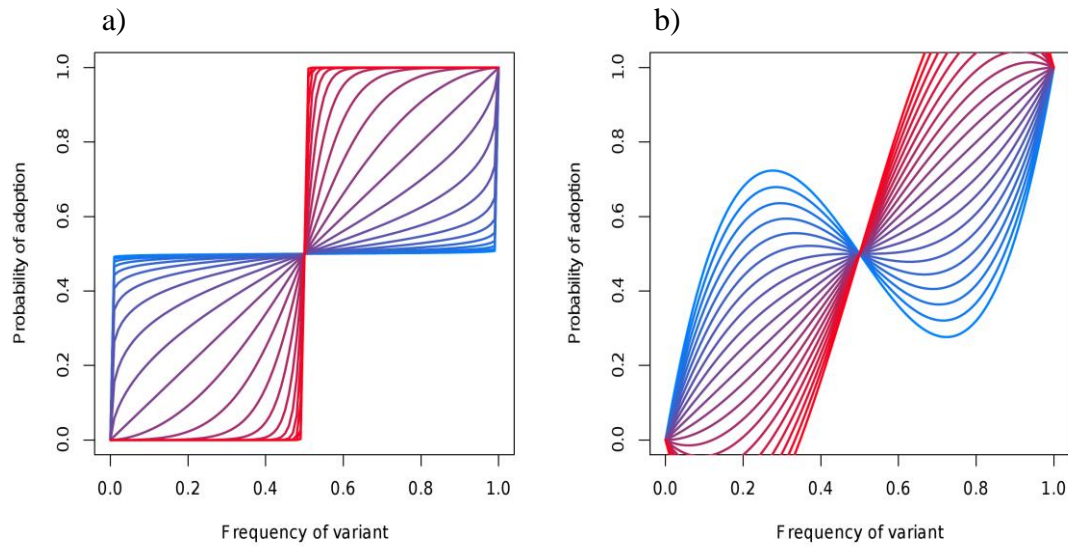


Figure 2.2

Two alternative forms of conformist transmission, where lines corresponding to conformist transmission are in red and anti-conformist transmission in blue. (a) Is calculated such that:

$$p = q^s / (q^s + (1-q)^s),$$

where  $p$  is the probability of adoption,  $q$  is the frequency of the variant and  $s$  ( $s \geq 0$ ) is the shape parameter that dictates whether the curve is conformist or anti-conformist. A benefit of this function is that it is constrained to produce plausible values. However, it does not allow for anti-conformist transmission strong enough to favour minority variants to the extent that they can become majority variants. (b) Shows the formulation used by Boyd & Richerson (1985) and Kendal et al. (2009) such that:

$$p = q + d * p * (1-p) * (2 * p - 1)$$

where  $p$  is the probability of adoption,  $q$  is the frequency of the variant and  $d$  dictates whether the curve is conformist or anti-conformist. This form can allow minority variants to be preferred to majority variants, but has the disadvantage that it is also capable of producing impossible adoption probabilities as shown by many of the red lines going outside the range 0-1.

Whilst conformist transmission has a single population-level stable state - behavioural homogenisation - anti-conformist transmission can have several (see **Fig. 2.3**). For example, in cases where the majority opinions are favoured over minority opinions, but to an insufficient extent to be conformist (i.e., the probability of choosing option

A in the hypothetical example, whilst  $<70\%$ , is  $>50\%$ , see **Fig. 2.2a**), the population will tend to a state of equal mixing with all options equally prevalent (see **Fig. 2.3a and b**). However, when minority opinions can be favoured over majority opinions (i.e., the probability of choosing option A in the hypothetical example is  $<50\%$ ), as has been considered theoretically (Kendal et al., 2009), several outcomes can result. A small bias towards the minority still results in equal mixing (see **Fig. 2.3c and d**). A stronger bias can result in looping oscillations where although fixation or equal mixing are metastable, dynamic switching between two mixed states is the only long-term outcome (see **Fig. 2.3e and f**). Finally, if the bias towards the minority is sufficiently powerful then chaotic dynamics can ensue (see **Fig. 2.3g and h**). As one of the strengths of conformist transmission was its specific population level outcome, the current grouping of these multiple different population level outcomes under the single term of anti-conformist transmission is ambiguous. Researchers would be well advised to bear in mind and make clear the population effects of any anti-conformist transmission they study.

Finally, I shall briefly consider transmission unbiased with regards to frequency. This refers to a form of frequency dependent learning where the probability of adopting a given trait is equal to its prevalence in the population. Unlike conformist and anti-conformist transmission this does not change the relative frequencies of variants in a population. It should be noted that this form of transmission is only unbiased with regards to frequency, being equivalent to an individual copying someone chosen at random, and it does not imply that all variants are equally likely to be chosen.

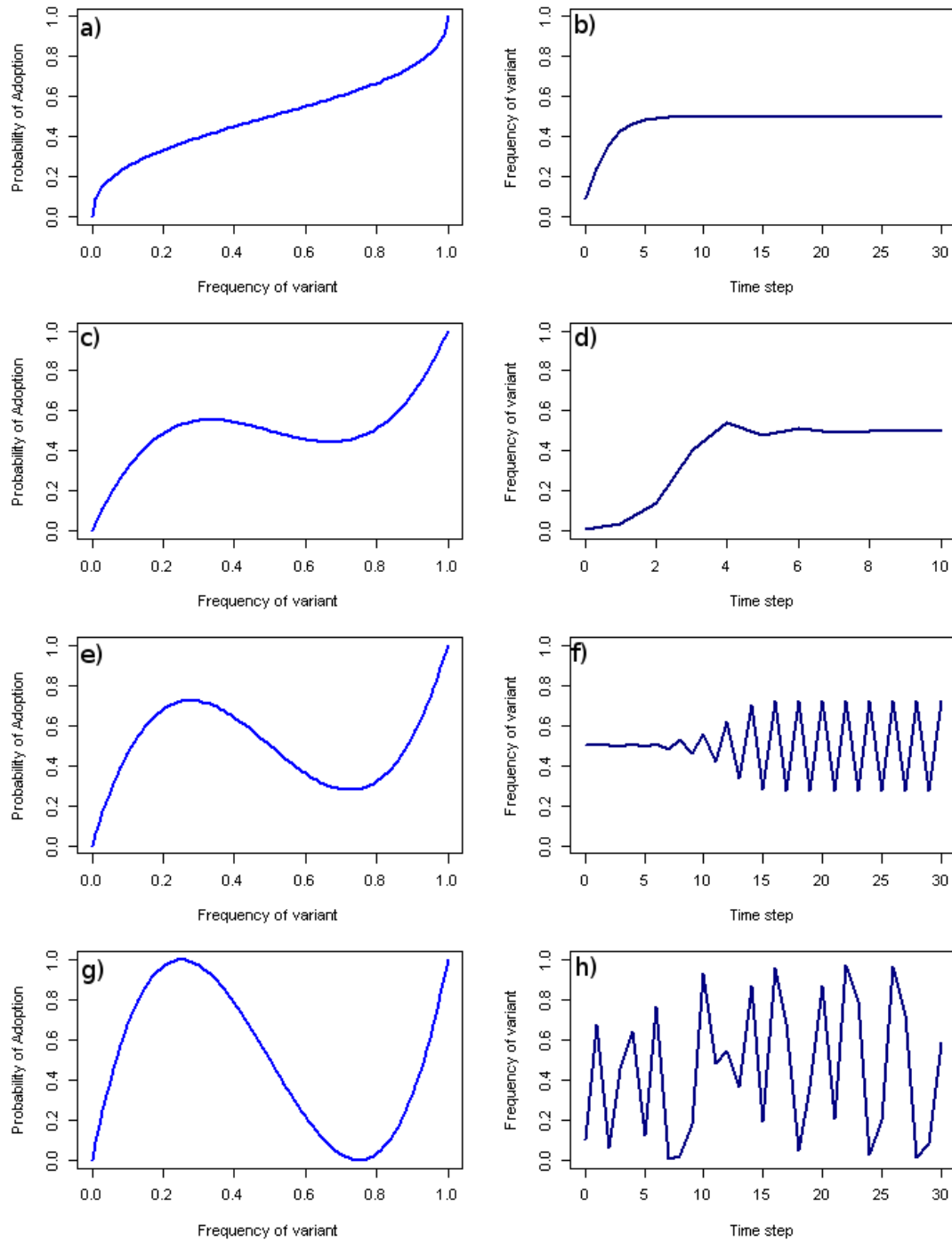


Figure 2.3

This figure shows the consequences of different forms and strengths of anti-conformist transmission. a) & b) use the formulation shown in figure 2.2 a) with  $s = 0.5$ . Panels c) - h) use the formulation shown in figure 2.2 b) with: c) & d)  $d = -3$ , e) & f)  $d = -5$ , g) and h)  $d = -8$ . The left hand panels (a, c, e and g) show the shape of the anti-conformist response to variant frequency. The right hand panels (b, d, f and h) show how an example population evolves over time in response to the anti-conformist bias. In the first two cases (a, b, c and d) a mixed population with equal variant frequencies is the stable outcome, although in panels c) and d) this is achieved following some small, damped oscillations that rapidly disappear. In panels e) and f) the stable outcome is to oscillate between two mixed states. Panel f) shows that this is achieved even if the population starts at a point of near equal mixing. In panels g) and h) chaotic dynamics ensue.

### **2.3 - Cultural evolution experiments**

Given the predictions made by evolutionary models concerning the success of conformist transmission as a learning rule, several experiments on adult human participants have been carried out with the required precision to distinguish a disproportionate tendency to adopt the majority decision from other rules that lack the same population level consequences. In the context of Tinbergen's four questions, these studies combine the study of function (investigating the adaptive value of conformist transmission) with the study of mechanism (investigating the contexts that elicit conformist transmission). Efferson, et al. (2008) carried out an experiment in which, over many rounds, participants repeatedly chose between two "technologies". The participants knew the alternative technologies had different expected payoffs, but did not know which was the better technology. Half of the participants were constrained to be asocial learners and were given feedback concerning the payoffs of their decisions, the other participants were constrained to be social learners and were only given information on the decisions of the asocial learners, but were not informed of their payoffs. Although conformist transmission was found to be an effective strategy for the social learners, Efferson et al. (2008) found that only the behaviour of some participants in the social learning condition, notably those that self-defined as conformist, could be well explained with a conformist model. The behaviour of the other participants, who did not describe their behaviour as conformist, could not. Efferson et al. (2008) characterise this difference in terms of a mixed population of conformists and "mavericks", the latter representing individuals typically reliant on asocial information. However, there was considerable individual-level variation within the conformist and maverick groups such that a dichotomy of types would not

be an appropriate interpretation. Rather, the data imply that individuals vary in the extent to which they utilise social information and/or are conformist.

In a similar experiment McElreath et al. (2005) used a two-armed bandit design where participants were required to choose between planting two types of crop 20 times at each of 6 different farms (totalling 120 decisions). The two crops were noisily associated with different payoffs and on some of the farms the payoffs could change over time. After each decision they could choose to see the decisions made by all other individuals playing at the same time as them (participants played in groups of 4-7 individuals). Although participants did show evidence a conformist response, they were better characterised by unbiased transmission when the environment was stable across time, a result at odds with theory that suggests environmental stability over time favours conformist transmission (Nakahashi et al., 2012). In addition to this, Toelch et al. (2010) found that subjects track variant popularity over time and in effect anticipate a future majority choice by favouring variants that show increasing popularity. This makes sense in the context of possible environmental change and the risks of the cultural inertia created by conformist transmission as such behaviour may allow individuals rapidly to take advantage of emerging technologies and overcome the cultural inertia that conformist transmission imposes.

It remains an outstanding question why the predictions of theoretical models - finding clear support for conformist transmission under many conditions - receive limited support from empirical data and this is a topic that I address experimentally in **chapter 3**.

### 2.3.1 - Social Psychological Studies

The huge amount of empirical data on conformity from social psychology (see **chapter 1, section 1.2**), which focuses on the social contexts that elicit social information use, might be thought to present an opportunity to clarify the issues raised by disagreements between models and experiments investigating conformist transmission. However, this is unfortunately not the case, for two reasons. Firstly, although a conformist placed in the Asch experimental paradigm would be expected to go along with the group decision, such experiments are unable to distinguish between multiple possible learning rules that posit a positive relationship between trait popularity and probability of trait adoption. For example, as depicted in **Figure 2.1**, conformist transmission, anti-conformist transmission and random copying all result in more popular traits being more likely to be adopted than less popular traits (Boyd and Richerson, 1985). Additionally several other learning strategies that are not directly affected by frequency can lead individuals to adopt decisions of the majority (Haun et al., 2012). For example, any given individual, including a prestigious individual, is more likely to be a member of the majority than the minority. Thus, an individual guided by prestige biased learning will be more likely to adopt the majority decision, not because of its majority status, but because prestigious individuals (along with all other individuals) are more likely to be members of the majority than the minority. By collecting data on which individuals are prestigious these different, non-frequency dependent mechanisms should be easy to identify; however, it is only through experiments that directly manipulate the level of consensus that we can distinguish conformist transmission from other frequency dependent learning rules. Secondly, a disproportionate tendency to adopt the majority behaviour is only

expected in cases where the observing individual is naïve (Boyd and Richerson, 1985). This means that the Asch paradigm is unsuitable to investigate conformist transmission as the exposure to, and simplicity of, the task used meant that the subjects were far from naïve when listening to the decisions of the confederates. Instead, asocial information must be controlled for, either experimentally, by using a design such that subjects genuinely are in a state of naïvety, or statistically, such that a measure of asocial information is taken and can be used in analyses to separate the effects of asocial and social information.

## **2.4 - Animal experimentation**

As cultural evolution takes an explicitly evolutionary approach, studies employing the comparative method are of value. Such studies provide further insights into the third of Tinbergen's questions, evolutionary history, as through a consideration of the current taxonomic distribution of conformist transmission researchers are potentially able to infer the most likely evolutionary history of the trait. In fact, evidence in line with conformist transmission exists in a wide range of taxa including fish (Day, 2001), rats (Galef and Whiskin, 2008; Konopasky and Telegdy, 1977), monkeys (Dindo et al., 2009) and great apes (Whiten et al., 2005), although it should be noted in the latter case the claim for conformity rests on a definition of conformity different to that from both social psychology and conformist transmission. Additionally, the methods employed in these studies, as in the Asch experiment, are typically not sufficient to rule out other forms of social learning that involve a positive relationship between trait popularity and the likelihood of its adoption. This is both because of the different goals of researchers, often interested in animal social learning more

generally than conformist transmission in particular, but also because of the logistical difficulties in determining something as specific as conformist transmission in an animal study. Only two published studies of which I am aware possess the precision to investigate conformist transmission directly. One is Pike and Laland's (2010) investigation of public information use in sticklebacks. In this case, although the behaviour in the paper is described as conformist, the sticklebacks' were found to be minimally sensitive to small majorities (e.g., 2v4), but showed a large increase in influence for larger majorities (e.g., 1v5). This pattern of behaviour is consistent with anti-conformist transmission. The second study is Haun et al.'s (2012) study of social learning in children, chimpanzees and orangutans. In this study, individuals could choose to put tokens in one of three boxes following demonstrations by three conspecifics, two of whom used the 'majority' box, one of the 'minority' box and none of them the 'other' box. Although the analysis failed to take into account individual differences and so results should be treated with caution, the observed data suggest chimpanzees were biased towards the majority box and away from the minority box (although they seemed indifferent to the other box). Given the problems with the analysis and the somewhat small sample size, this experiment can only provide evidence consistent with conformist transmission and is not able to rule out alternatives. Additionally, the inclusion of a 3rd option muddies the water as there is little theoretical work available concerning conformist transmission on a 3-option task. Despite these problems the work does raise the valuable distinction between the number of individuals and the number of observations and in a second study attempts to tease the two apart. However, given the problems these issues are still far from resolved and more detailed experimental work is required to understand both the



evolutionary history of the human capability for conformist transmission and its phylogenetic distribution.

## **2.5 - The development of conformist transmission**

The fourth of Tinbergen's questions, ontogeny, is one area that the study of conformist transmission has left relatively untouched. However, multiple sources provide evidence consistent with a conformist bias in young child. For example, studies with preschool children have shown that they are more likely to rely on information from a group of adults in apparent consensus than from a single adult. When presented with conflicting names for a novel object by two different informants, 4-year-olds are receptive to the reactions of bystanders. If two bystanders signal agreement (via head nods and smiles) with one name and disagreement (via head shakes and frowns) with the other name, children overwhelmingly endorse the name eliciting bystander agreement. Similarly, if three informants all point to the same object as the referent of a novel name whereas a single informant points to a different object, 3- and 4-year-olds select the former when asked to identify the named object (Corriveau et al., 2009a). Three and 4-year-olds also defer to a consensus of three informants who produce judgments that conflict with what children can see for themselves. Having correctly identified the biggest line of a trio, children will change their judgment following a different claim by the consensus at rates similar to those observed in classic studies of conformity in adults. Thus, children often endorse a consensus when they lack relevant perceptual cues (as in learning names for novel objects) but they will even do so despite the availability of perceptual cues. Were such studies extended to examine the impact of different levels of

consensus it would be highly illustrative with regards to the ontogeny of conformity. Indeed, there is already evidence that learning rules vary over time. For example, children of different ages show a shift in sensitivity to reliability assessment that may be experientially triggered (Clement et al., 2004).

## **2.6 – The neurobiology of conformist transmission**

Whilst there are a small number of studies investigating the neurobiology of social information use (see **Chapter 1, section 1.7**), there are no studies that have directly investigated the neurobiology of conformist transmission. However, there are some neurological studies that are of relevance. For subjects to be employing a conformist learning bias it might be expected that parts of the brain would be identified that evaluate levels of consensus amongst demonstrators. Whilst no data exists for studies using sufficiently large groups of demonstrators with varying levels of consensus, there are nonetheless hints of the existence of such a mechanism in the brain. A music choice experiment (Campbell-Meiklejohn et al., 2010) found that along with activity in the insula cortex and right tempoparietal junction, areas associated with monitoring the decisions of others, anterior insula activity increased when the two “expert” reviewers were in agreement. Although this is suggestive of a consensus evaluating mechanism it should be noted that with a group of only 2 demonstrators the social information was either in unanimity or in total disagreement, thus the anterior insula may have been responding to social information with an overall message and not the specific level of consensus itself. What is clearer, however, are changes in brain activity caused by disagreement between the subject and the demonstrators. Klucharev et al. (2009) found that disagreement between the subject and the

demonstrators caused activity in several areas known to be involved with more general error and conflict processing such as the rostral cingulate zone (Botvinick et al., 2004) and depressed activity in reward centres such as the nucleus accumbens in the ventral striatum. Thus, brain areas that evaluate object value, such as the ventral striatum in the music choice task (Campbell-Meiklejohn et al., 2010), also seem to play a role in rewarding the subject for being in agreement with others. Furthermore the magnitude of the change in activation of these areas predicted changes in subsequent subject behaviour (Campbell-Meiklejohn et al., 2010; Klucharev et al., 2009).

## **2.7 - Conclusions**

From the above it can be seen that Tinbergen's four questions of history, ontogeny, function and causation are effective in identifying areas in which current understanding of conformist transmission is well developed, and others where it still needs work. With large amounts of theoretical and empirical data on the topic, researchers are beginning to identify the circumstances under which individuals will conform, and with further careful non-human experimental work they will soon be able to understand the current taxonomic distribution of such a bias. However, there is still much work to be done on understanding the contexts in which conformist transmission is displayed by humans and the findings of theory and experiment need to be reconciled.

In the next chapter I present the results of a series of experiments with adult humans investigating a range of social learning strategies, including conformist transmission.

In particular the results shed light on the specifics of payoff biased learning and conformist transmission and also aim to test the prediction that social learning strategies should be adaptive in terms of their effect on the ability to make accurate decisions. In **chapter 4** I present a similar experiment with young children investigating the development of conformist transmission and a bias to copy when uncertain. Again I seek to understand the extent to which the observed behaviour satisfies evolutionary predictions of adaptiveness. In **chapter 5** I present the results of an experiment investigating the role of social learning in the evolution of hominin stone tool technology. Whilst not as general a study as those described in the preceeding chapters, both human tool use and culture are uniquely complex and so an investigation of the relationship between the two is of great interest for understanding human evolution.

## Chapter 3

# **An Empirical Study of Social Learning Strategies in Adults**

## **Abstract**

As outlined in Chapters 1 and 2, successful decision making requires the complex integration of social and asocial information to generate effective learning and decision making. Recent formal theory predicts that natural selection should favour adaptive learning strategies, but relevant empirical work is scarce and rarely examines multiple strategies or tasks. In this chapter I present a test of nine hypotheses derived from theoretical models, through four experiments investigating factors affecting when and how humans use social information, and whether such behaviour is adaptive, across several computer-based tasks. The number of demonstrators, consensus among demonstrators, confidence of subjects, task difficulty, number of sessions, cost of asocial learning, subject performance and demonstrator performance all influenced subjects' use of social information, and did so adaptively. The analysis provides strong support for the hypothesis that human social learning is regulated by adaptive learning rules.

### 3.1 – Introduction

Despite the large body of theoretical and empirical work described in the previous two chapters there are still several outstanding questions concerning adult human social learning that need to be addressed. In particular, the existence of conformist transmission (see **Chapter 2**) is far from clear. Whilst theoretical work has generally found conformist transmission to be a successful strategy, this has not been replicated empirically and more detailed data is needed to evaluate this apparent discrepancy. Additionally, the nature of payoff biased social learning (see **Chapter 1, section 1.8**) is also under question. Whilst it is favoured theoretically and supported empirically, there are several forms that such a rule could take (Schlag, 1998) and as yet there is not data capable of discerning between these different possibilities.

Beyond these questions there are also some consistent limitations of previous work. For example, previous research into social learning strategies has typically made the implicit simplifying assumption that individuals only employ one strategy in a given context (Laland, 2004), with models rarely considering more than one strategy, and experiments designed in such a way that subjects cannot use multiple strategies simultaneously. For example, in an arrow-head design task (Mesoudi and O'Brien, 2008) individuals were not given free access to social information, but instead clicked mutually exclusive buttons corresponding to solely frequency guided or solely payoff-guided social learning, preventing the use of multiple social learning strategies simultaneously. Although this may be a reasonable step when investigating the existence of a particular strategy, it prevents consideration of interactions between

factors and prevents a full understanding of social learning. Another limitation is that empirical work is typically limited to a single experimental task. The result being, that although social learning biases are typically described as general features there is as yet very little data with which to directly evaluate this hypothesis. Although evolutionary theory suggests these rules should be broadly adaptive, and thus be in use across a range of tasks, supporting this simplification, it still requires testing. A final limitation is the paucity of empirical data on whether the strategies individuals use are employed adaptively in the sense that they aid successful decision making (a notable exception being Henrich and Henrich, 2010). Most studies, being among the first of the kind, have aimed simply to document the existence of such strategies. This, however, ignores a central prediction of the cultural evolutionary models - that such learning biases should be adaptive. Indeed, were such learning biases to be documented, and their effects to be observed to be generally maladaptive, this would be as much a problem for the theoretical literature as for the biases to not be observed at all.

In this chapter I present data collected across 4 experiments in which adult human participants were required to solve computer-based tasks by utilising asocial and/or social information. In total, 6 diverse tasks and 2 methods of information presentation were used and the effects of 9 different factors were investigated. Factors were often allowed to vary simultaneously, allowing investigation of interactions on participant reliance on social information use. The inclusion of each factor can be considered as a test of a specific hypothesis generated by the theoretical cultural evolution literature (Boyd and Richerson, 1985; Henrich and Boyd, 1998; Schlag, 1998; Boyd and



Richerson, 1988). Factors considered are the number of demonstrators (+), demonstrator consensus (+), subject confidence (-), task difficulty (+), the cost of asocial learning (+), task familiarity (-), demonstrator (+) and subject (-) performance, and their difference (+), where + and - signs indicate a prediction for a positive or negative relationship with social learning. Additionally, the use of multiple tasks allows a test of the previously implicit assumption that social learning strategies are deployed consistently across diverse contexts and sensory modalities, whilst investigating numerous factors allows a test of whether individuals are limited to using a single strategy at a time or if such factors can interact. Furthermore, by setting social information use in a task-solving context we are able to investigate the adaptiveness of participants' behaviour by measuring participants' success in the tasks. Thus, the work presented here aims to provide a comprehensive empirical test of the adaptive learning strategies predicted by evolutionary models, to provide new data particularly concerning current areas of debate such as pay-off and frequency dependent social learning and to extend existing work by examining the simultaneous use of multiple strategies.

## **3.2 - Methods**

### **3.2.1 - General Methods**

I carried out a series of 4 computer-based experiments, the first of which was subsequently repeated under differing conditions. In each of the experiments, participants were presented with multiple trials of one or more binary-choice tasks that they had to solve using asocial and/or social information. By asocial information I

mean access to the task itself, whereas social information provides information concerning the decisions of other individuals (henceforth "demonstrators") who themselves had had direct access to the task. Experimental sessions lasted for 60 minutes. Participants were paid £5 for taking part, plus a bonus of up to £5 dependent on their performance during the experiment.

### **3.2.1.1 - Participants and Apparatus**

In total, 99 participants took part in the first experiment (45 in the first instance and a further 54 when it was repeated), 57 in experiment 2, 38 in experiment 3 and 61 in experiment 4. Participants' ages ranged from 18 to 58 (mean=22, s.d.=4); 83 were male, 172 were female and all but 9 were students of the University of St Andrews.

Participants took part in batches of 1-11 individuals in the Media Lab, of the Bute Medical Building of the University of St Andrews. All participants had access to a computer and were separated by large screens such that they could not see each other. Participants were provided with headphones such that they could listen to any required sounds without disturbing each other.

### **3.2.1.2 - Procedure**

On arrival, the experimenter read participants a briefing and gave them the opportunity to ask any questions. The experimenter then went to an adjoining room and remotely started the experiment program on the participants' computers. The connecting door was kept open so that the experimenter could hear if any participants communicated, but no instances of between-participant communication were observed. When all participants had completed the experiment the experimenter returned to debrief the participants who were then free to collect payment and leave.

### **3.2.1.3 - General Design**

The four experiments were designed to repeatedly investigate the effect of numerous factors on participants' social learning in a range of contexts and involved two methods of information presentation, 6 tasks and three sources of social information. I first describe these general design components and then describe how each of the four experiments was constructed from these elements. Each experiment was designed without knowledge of the results of any others. Thus they can be viewed as independent attempts to address a set of related questions about human social learning. As most factors were investigated in more than one of the experiments, and although the results for each experiment are presented separately, I discuss the results

together, rather than each in turn, so as to avoid repetition and to consider the results as a whole.

#### **3.2.1.4 - Methods of Information Presentation**

Across the four experiments two methods of information presentation were used. The first I shall refer to as the linear protocol. Under this protocol participants received asocial information first, and then social information. After receiving each type of information participants were required to make a decision and rate their confidence in their decision on a 7-point Likert scale (Likert, 1932). With this protocol, only in cases where a participant's initial decision disagreed with the modal decision of the demonstrators they subsequently saw the decisions of was is clear as to whether a participant's final decision was based on asocial or social information. Accordingly, participants were considered to have used social information if their final decision (following social information) differed from their initial decision (following asocial information), given that the majority of demonstrators disagreed with the participants' initial decision. In experiments using this protocol access to the task (i.e. asocial information) was unavoidable and of a fixed duration, accordingly, participant's incurred no cost to access the asocial information. If such a cost were implemented it would be same for all participants across all trials and so be equivalent to the scenario of no cost. Given this the asocial information can be considered "free".

In the second protocol, termed the parallel protocol, participants chose between either social or asocial information. Whichever information source they chose, they did not receive any information from the other source, i.e., a participant choosing social information would receive no asocial information and so would not have direct access to the task on that trial. Before making this decision participants were informed of some specifics of the trial, the nature of which differed across experiments and are detailed below (**Section 3.2.4**). After receiving their chosen source of information participants were required to make a single decision. Using this protocol it is always clear which information source participants are using and participants were considered to have used social information when they chose to view social over asocial information. Under this protocol direct access to the task (i.e., asocial information) is optional and participants were able to determine the duration or number of presentations of the task stimulus. In this case a cost was levied on asocial information use that was directly proportional to number of presentations or presentation duration.

### **3.2.2 - The Tasks**

#### **3.2.2.1 - The foraging task**

This task involved selecting the more profitable of two virtual foraging sites and was designed to mimic the procedures deployed in published animal experiments (Coolen

et al., 2003, 2005; van Bergen et al., 2004). Participants were shown two possible foraging sites, represented as squares side-by-side on the participants' screen. Participants subsequently received 10 visual presentations of a number of apples on each of the two sites, totalling 20 presentations in all. The true values of the sites were either 8 and 4, or 6 and 2, with which site was the most profitable randomised across trials. The number of apples shown in each presentation was independently drawn from a Poisson distribution with a mean equal to the true distribution of the site. However, to add further uncertainty, for one of the ten apple presentations at each site the number of apples shown was drawn from the distribution of the other site. The apple presentations were given in a random order and each lasted 500ms. The location of each apple on its site was randomised for each apple and for each presentation.

### **3.2.2.2 - The mental rotation task**

Participants were required to decide whether two shapes were the same shape seen from different angles or different shapes entirely (see **Fig. 3.1a**). In each trial, participants received a single visual presentation of an image of two shapes, the duration of which varied across experiments. The shapes used were constructed from 10 identical cubes arranged into 4 orthogonal arms of 3, 3, 4 and 3 cubes. Where shapes were different, the second shape was a reflection of the first. Whether the same or different, pairs of shapes were always rotated relative to one another by between 20 and 160 degrees. This task was based upon on that used by Shepard & Metzler (1971),

and allows trials of different difficulty to be generated as trial difficulty is related to the angle by which shapes have been rotated relative to each other.

### **3.2.2.3 - The length-estimation task**

Participants were required to decide which of two irregular lines was longest (see Fig. 3.1b). In each trial, participants received a single visual presentation of an image of two lines, the duration of which varied as detailed below. This task was chosen as it is hard to solve through visual observation alone and so all responses will be associated with some degree of uncertainty.

### **3.2.2.4 - Audio tasks**

Three different audio tasks were used. Across the three, participants were required to decide which of two tones was (i) higher in pitch (the “pitch task”), (ii) greater in intensity (the “intensity task”) or (iii) whether a single tone was increasing or decreasing in pitch (the “pitch-modulation task”). These tasks were chosen to provide a test of the sensory inter-modality of learning strategies as the other tasks primarily involved the visual sensory system which was not at all involved in the audio tasks. In trials of all three audio tasks, participants had access to button(s) that played the relevant tone(s) and could play them as many times as they wished, but incurred a

cost per playing, detailed below (see **Section 3.2.4.4**), that reduced their bonus payment.

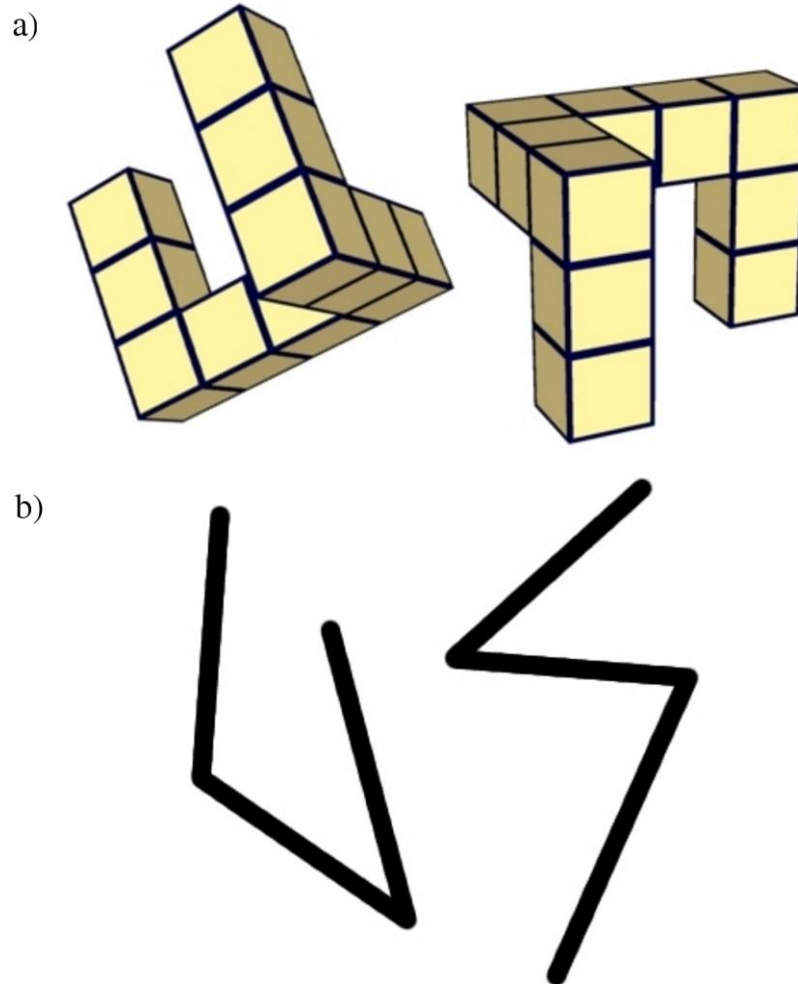


Figure 3.1

(a) An example image of a pair of shapes used in the mental rotation task. In this case the two shapes match in that they are the same shapes seen from two different angles. (b) An example image of a pair of lines used in the length estimation task, in this case the right line is the longer of the two.



### 3.2.3 - Types of Social Information

The decisions of demonstrator individuals were presented using two tiles labelled with the possible answers (e.g., "the shapes match" and "the shapes do not match" in the case of the mental rotation task). A single demonstrator's decision was represented by the relevant tile flashing red for 350ms. This information was either from a battery of previous participant responses (battery), from other participants taking part in the same experimental session (live), or was generated by me (manipulated). When social information was manipulated I employed a conditional information lottery (Bardsley, 2000) to ensure participants would treat the information as genuine without being deceived. Under this procedure, whilst most trials used manipulated social information, a minority of trials used genuine (battery) social information. Participants were not told which trials used which type of social information, but were told that only on trials using the genuine social information would their performance be scored and affect their payment. Under this protocol, if a trial uses genuine information the participant benefits from trusting it as it will help their decision making, but suffers a cost if they do not trust it as they do not gain the benefit. If a trial uses manipulated social information, in which case participant decisions do not affect their payment, the participant neither benefits nor suffers a cost from trusting the social information or not doing so. Accordingly, and given that participants cannot tell the source of social information on a particular trial, the optimal strategy for participants is to trust all social information. Supporting this suggestion, it has been shown that under the conditional information lottery, participants will treat all information as genuine (Bardsley, 2000). Whenever the conditional information

lottery was used, participant feedback, collected following the experiment, was used to identify participants who had failed to understand this procedure and their data was excluded from analysis.

### **3.2.4 - The experiments**

#### **3.2.4.1 - Experiment 3.1**

Participants completed 10 trials of the foraging task, using the linear protocol and manipulated social information, presented in a random order. I manipulated (i) the number of demonstrators and (ii) the level of consensus in the social information they provided such that social information consisted of either 4 v. 0, 5 v. 3, 6 v. 2, 7 v. 5 or 8 v. 4 demonstrators. Three of these levels (4v0, 6v2, 8v4) were chosen to replicate social conditions used in previous work with sticklebacks and with the intention to carry out cross-species comparisons; however, these comparisons were later abandoned and other levels were added to give the social information added realism. This experiment was subsequently repeated, with the only difference being that the order of information presentation was reversed (i.e. social information was presented first). This was in order to investigate the possibility that participants were differentially influenced by information according to the order in which it was presented.

### **3.2.4.2 - Experiment 3.2**

Subjects completed 24 trials of the mental-rotation task, using the linear protocol and manipulated social information, presented in a random order. The shape images were presented for a fixed time of 4s. I manipulated (i) the number of demonstrators and (ii) the level of consensus in the social information as in Experiment 1 (see Section 3.2.4.1).

### **3.2.4.3 - Experiment 3.3**

Participants completed 60 trials using the parallel protocol and battery social information. The trials were arranged into 4 groups of 15, the 1st and 3<sup>rd</sup> groups used the mental-rotation task and the 2<sup>nd</sup> and 4<sup>th</sup>, the length-estimation task. This structure was adopted to avoid participants becoming bored and not attending to trials. The order of trials within each group was randomised for each participant.

Prior to choosing between the two sources of information (as per the parallel protocol), participants were informed of question difficulty ('very easy', 'easy', 'moderate', 'hard', or 'very hard') and the number of demonstrators they would see the decisions of should they choose social information (1, 3 or 5). For the mental-rotation task difficulty was calculated as the proportion of previous participants that

got the question wrong following asocial information. Previous work has suggested that the angle of rotation predicts trial difficulty; however, I found this not to be the case. For the length-estimation task difficulty was negatively related to the Weber fraction of line length, such that:

$$\text{difficulty} \propto -(\text{length of longer line} / \text{length of shorter line}),$$

as this is typically a reliable measure of difficulty for perceptual tasks. Participants who chose asocial information were able to examine the relevant image indefinitely, but at a cost proportional to the time spent observing that reduced their possible bonus payment. Participants were informed of the cost incurred after each trial. Social information was free.

#### **3.2.4.4 - Experiment 3.4**

Subjects completed 45 trials using the parallel protocol arranged into 3 groups of 15, that deployed the pitch task, intensity task and pitch modulation task respectively. The first two trial groups used battery social information and within each group trial order was randomised for each participant. The third trial group used live social information requiring participants to complete individual trials simultaneously such that trial order could not be randomised between participants, trial order was randomised between

batches of participants however. If a participant chose to copy an individual who was copying someone themselves, the participant received only a message informing them of this, and therefore had to guess.

Prior to choosing a source of information, in trials using the pitch task participants were informed of the number of demonstrators available (1, 3 or 5) and the cost incurred each time they played a tone (10, 20 or 30% of maximum pay-off). In trials using the intensity task, the social information was always from a group of four demonstrators and participants were informed of the degree (but not direction) of consensus amongst the demonstrators (0, 0.5 or 1), whilst the cost per playing of the tones was held constant and participants were informed at the end of each trial of the cost incurred. In the pitch-modulation task participants were ranked on their performance, assayed by summing their pay-off across all pitch-modulation trials irrespective of the source of information participants had used, and were informed of their own rank and the rank of their single potential demonstrator at the start of each trial.

### **3.3 - Analyses**

Unless otherwise stated, I analysed the data using Generalised Linear Mixed Models (GLMMs) with a Bernoulli error structure and logit link function. I used Markov Chain Monte Carlo (MCMC) methods (Bolker et al., 2009; Kruschke, 2010) to fit the models in WinBUGS 1.4 (Lunn and Spiegelhalter, 2009) and to generate credible intervals for each parameter. Minimal adequate models were constructed by

backwards elimination, removing variables for which the 90% central credible interval included 0 from a model containing all predictors and second-order interactions in addition to baseline values and individual level effects. Parameter values were estimated using a sample of at least 3000 iterations, after a suitable burn-in period and thinning to remove autocorrelations. Preliminary analyses were carried out to establish whether factors should be modelled as linear or categorical.

Unless otherwise stated, all figures show the median of posterior samples from the fitted models, as estimates of parameter values. Plots of the posterior parameter estimates are preferred to raw data because they directly illustrate the marginal effect sizes associated with a given predictor while controlling for the effects of other predictors. See **Figures 3.2a) and b)** for a comparison of the raw data and model estimates.

### **3.3.1 - Experiment 3.1**

I modelled the probability that a participant whose initial asocial decision disagreed with the majority of demonstrators in the social information they subsequently received would switch their decision as a function of demonstrator number, demonstrator consensus (the proportion of demonstrators choosing the majority option, scaled to range between 0 and 1), and participant confidence in their initial decision. The final model contained a baseline level of copying, a random individual effect and the effects of participant confidence, demonstrator number and consensus, but not interactions.

To investigate whether the use of social information was adaptive I separately modelled the probability a participants' initial, asocial, decision was correct as a function of participant confidence. To investigate whether the order of information presentation affected participant reliance on the two sources of information, I modelled the probability a participant favoured social information over asocial as a function of order of information presentation. In the case of participants who received social information first, it is hard to discriminate trials where the two sources of information disagreed, but participants stuck with the social information from trials where the two sources of information agreed. To get around this I assumed that asocial information was never misleading, in that if the majority of demonstrators were incorrect then the asocial information would always disagree with the social information.

### **3.3.2 - Experiment 3.2**

I modelled the probability that a participant whose initial asocial decision disagreed with the majority of demonstrators in the social information they subsequently received would switch their decision as a function of the number of demonstrators, demonstrator consensus, and participant confidence in their initial decision. The final model contained a baseline level of copying, a random individual effect and the effects of demonstrator consensus and participant confidence, but no interactions.

To investigate whether the use of social information was adaptive I separately modelled the probability a participants' initial, asocial, decision was correct as a function of participant confidence. To further investigate the nature of participants' response to consensus (i.e., whether it was conformist), two further analyses were carried out. Neither of these analyses required that data be limited to cases where the social information disagreed with the participant's initial, asocial, decision. Firstly, I modelled the probability a participant would change their decision as a function of demonstrator number, the proportion of demonstrators that disagreed with the participant, participant confidence in their initial decision. Secondly I modelled the probability that a participant's final decision would be that the shapes matched as a function of the number of demonstrators, the proportion of demonstrators deciding the shapes matched (which was separately modelled as both linear and categorical to avoid constraining the shape of the response), whether the participant initially decided the shapes matched, participant confidence in their initial decision, whether the shapes actually matched or not, and random trial effects. Interactions between these effects were included when proportion was modelled linearly, but when proportion was modelled as categorical interactions between proportion and the other effects could not be included.

### **3.3.3 - Experiment 3.3**

I modelled the probability that an individual would choose to view social information instead of asocial information as a function of demonstrator number, task difficulty, session number, and previous costs incurred using asocial information. The final model contained a baseline level of social information use, a random individual effect



and the effects of trial difficulty, previous costs incurred, number of demonstrators, session and the interaction between demonstrator number and session.

To investigate whether the use of social information was adaptive I modelled the pay-off to participants on trials where they used asocial information (using a normal error structure and without a link function) as a function of task difficulty, trial group, and task.

### **3.3.4 - Experiment 3.4**

For each of the three tasks I modelled the probability that an individual would choose social information instead of asocial information. In the pitch task this was as a function of demonstrator number and the cost to play the sounds, in the intensity task as a function of demonstrator consensus and costs incurred when using asocial information previously, and in the pitch-modulation task as a function of participant rank, demonstrator rank and the difference between the two. The final model for each of the three tasks contained a baseline level of social information use and a random individual effect. The final model for the “pitch task” also included the effect of demonstrator number and the cost to listen to the sounds. The “intensity task” final model contained the effect of demonstrator consensus whilst the “pitch modulation task” final model also included demonstrator rank, the difference in rank between participant and demonstrator and the interaction between these two.

To investigate whether the use of social information was adaptive I modelled the probability that a participant would make the correct decision on a trial for which they had chosen asocial information as a function of participant performance ranking.

### 3.4 - Results

#### 3.4.1 - Experiment 3.1

Participants were increasingly likely to adopt a conflicting majority decision with decreasing confidence in their own decision (gradient of slope on the linear predictor: median = -0.48, 95% central credible interval: [-0.87, -0.13], **Fig. 3.2e**), increasing demonstrator consensus (2.50, [1.23, 4.01], **Fig. 3.2c**), and increasing numbers of demonstrators (0.11, [0.00, 0.23], **Fig. 3.2d**). There were no interactions. Participant confidence predicted whether a participant was right or wrong, with low confidence predicting an incorrect answer (0.50, [0.19, 0.84], **Fig. 3.2e**). There was no effect of information presentation order (social/asocial first) on social information use (0.01, [-0.50, 0.49]). The minimal adequate model contained a baseline level of copying, a random subject effect and the effects of subject confidence, demonstrator number and consensus.

#### 3.4.2 - Experiment 3.2

As in Experiment 1, participants were increasingly likely to switch to a conflicting majority decision as demonstrator consensus increased (gradient of slope on the linear

predictor: median = 2.51, 95% central credible interval: [1.63, 3.41], **Fig. 3.2c**) and with decreasing confidence in their initial choice (-0.81, [-1.09, -0.56], **Fig. 3.2e**). However, there was no evidence for an effect of the number of demonstrators (-0.02, [-0.21, 0.14], **Fig. 3.2d**). Low participant confidence again predicted an incorrect answer (0.23, [0.11, 0.34], **Fig. 3.2e**). The minimal adequate model contained a baseline level of copying, a random subject effect and the effects of demonstrator consensus and subject confidence.

The first conformist transmission analysis showed that participants were increasingly likely to change their initial decision as the proportion of demonstrators that disagreed with them increased (5.12, [4.04, 6.44]). There was an interaction between demonstrator number and disagreement (0.44, [0.17, 0.76]), producing an S-shaped curve when demonstrator number was high and participant confidence was low (**Fig. 3.3a, b**).

The second conformist transmission analysis found that the likelihood a participant's final decision was that the shapes matched co-varied with the proportion of demonstrators reporting that the shapes matched (8.30, [6.56, 10.45], **Fig. 3.3d**), an effect that increased with the number of demonstrators (i.e. a proportion x number interaction, 0.66, [0.34, 1.02], 'proportion' modelled as linear). There was another interaction between a participant's initial decision and their confidence in that decision (initial decision: 5.90, [6.80, 7.90]; confidence: -0.13, [-0.27, 0.01]; interaction 0.69, [0.42, 0.96]). Participants were also more likely to decide the shapes

matched if they genuinely did (0.43, [0.03, 0.85]; if not: -0.44, [-0.77, -0.13]). When proportion was modelled as categorical the results were qualitatively the same, but a moderate consensus amongst the demonstrators had a disproportionate impact (**Fig. 3.3c**).

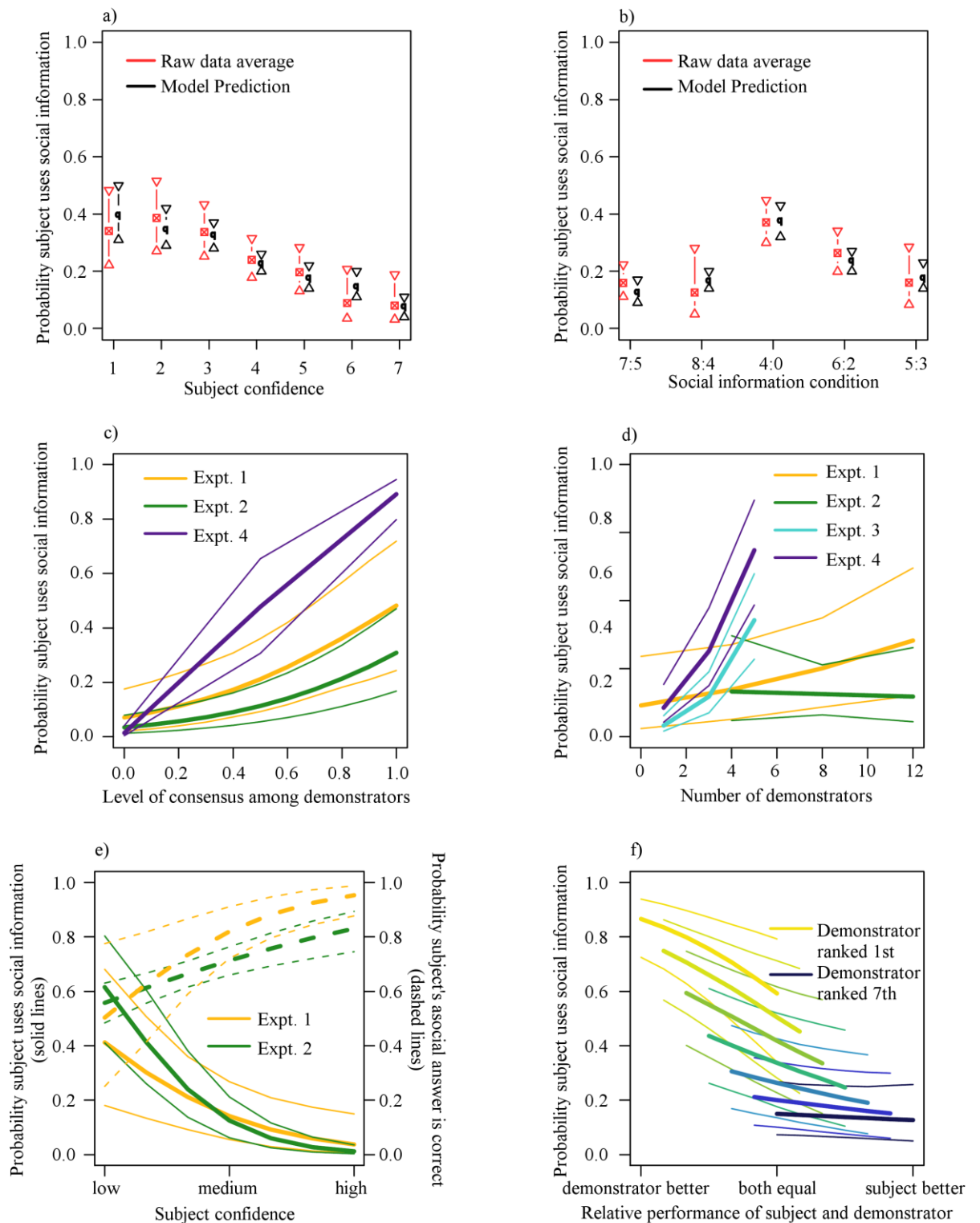


Figure 3.2

(a) & (b) A comparison of the raw data (red lines, showing the average value and Wilson confidence interval) and model predictions (black lines, showing the median and 95% central credible interval) for the probability of social information use given (a) subject confidence and (b) the different social information conditions (where 7:5 means 7 demonstrators make one decision whilst 5 make the other). The similarity between the two sets of lines show how

the model closely matches the raw data, whilst the narrower intervals for model estimates highlight the increase in precision offered by the model. (c) Median (thick lines) and 95% credible intervals (thin lines) for the probability that a subject uses social information depending on demonstrator consensus in Experiments 3.1 & 3.2 (modelled as linear) & 4 (modelled as categorical). (d) The probability that subjects use social information increases with the number of demonstrators in three of four experiments, with more pronounced effects in experiments 3.3 and 3.4, which used the parallel protocol. (e) In Experiments 3.1 & 3.2 the probability that a participant uses social information decreases with their confidence in their own judgement based on asocial information (solid lines), whilst the probability that the participant's initial decision is correct increases (dashed lines). (f) The probability a participant copies a demonstrator depends on both their relative performance and the absolute performance of the demonstrator. (Expt. 3.4). Participants respond differently to high performing demonstrators dependent on their own success, however, all participants are equally unlikely to copy a poorly performing demonstrator.

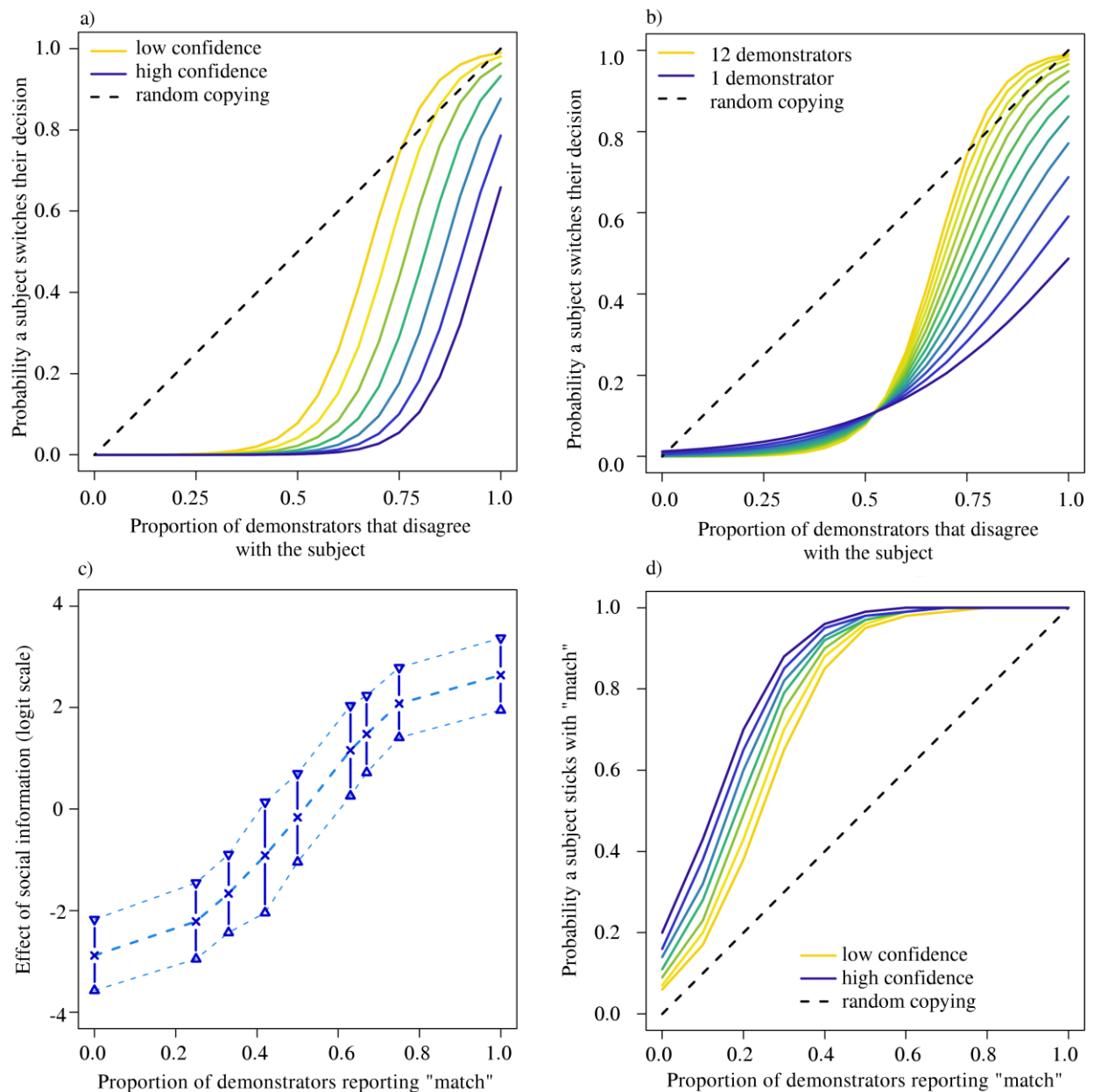


Figure 3.3

(a) and (b). The proportion of demonstrators that disagree with the participant's initial, asocial, choice strongly affects the probability that subjects will change their decision. (a) Decreases in participant's confidence in their initial, asocial decision (shown for the case where number of demonstrators = 12), and (b) increases in demonstrator number (shown for the case where participant confidence = low), increase the likelihood of a switch following conflicting social information. Participant behaviour, when uncertain and given many demonstrators, is consistent with conformist transmission in that participants are disproportionately likely to switch their decision when faced with a large opposing majority. The black dashed lines portray the expected result of random copying. (c) and (d). The proportion of demonstrators reporting the shapes match strongly affects the probability that a participant's final decision will be that the shapes match. (c) In line with a conformist response to social information, when unconstrained (i.e. modelled as categorical),

intermediate levels of consensus have a disproportionately large effect on decision making. The y-axis is the change to the linear predictor of the model, on which a change of magnitude 4 could alter the probability a participant decides the shapes match by as much as 76%. Accordingly, without other influences, social information is likely to have a dramatic effect on participant behaviour. (d) Participant behaviour, however, is also strongly affected by prior information and their confidence in it (the lines shown are for the case where participants already believe that the shapes match). Thus, although participant behaviour may not consistently be conformist, their response to the social information alone was conformist. The black dashed line portrays the expected result of random copying.

### 3.4.3 - Experiment 3.3

Subjects were more likely to choose to view social information on harder than easier trials (gradient of slope on the linear predictor: median = 0.35, 95% central credible interval: [0.25, 0.44]) and when they had incurred greater rather than lesser costs on the previous trial (0.077, [0.021, 0.14]). There was no evidence for an effect of task (-0.13, [-0.44, 0.18]). There was an interaction between the trial group and the number of demonstrators (demonstrator number: 0.55, [0.46, 0.63], **Fig. 3.2d**, trial group: -0.18, [-0.31, -0.06], interaction: 0.10, [0.03, 0.17]). Whilst participants increasingly chose social information as the number of demonstrators increased, over groups of trials participants became less likely to choose social information when there was only one (-0.37, [-0.60, -0.15]) or three (-0.18, [-0.30, -0.05]) demonstrators, but not when there were 5 (0.02, [-0.14, 0.16]). The minimal adequate model contained a baseline level of social information use, a random subject effect and the effects of difficulty, previous costs incurred, demonstrator number, session and the interaction between demonstrator number and session.



Subject pay-off when using asocial information increased over groups of trials (0.61, [0.42, 0.79]) and was negatively related with question difficulty in trials using the length estimation task (-0.20, [-0.39, -0.02]), but not those using the mental rotation task (0.14, [-0.04, 0.33]).

#### 3.4.4 - Experiment 3.4

The likelihood of participants choosing social information covaried positively with the consensus among demonstrators (modelled as categorical, as preliminary analyses suggested a linear model was inappropriate; 2v2: 0; 3v1: 4.13, [3.25, 5.18]; 4v0, 6.33, [5.34, 7.48], **Fig. 3.2c**), the number of demonstrators (gradient of slope on the linear predictor: median = 0.64, 95% central credible interval: [0.52, 0.78], **Fig. 3.2d**) and the cost to play the sounds (1.03, [0.79, 1.28]). There was no evidence that costs incurred on the previous question had no effect (0.11, [-0.19, 0.42]) or that participant rank had an effect (0.06, [-0.04, 0.15]). There was a positive interaction between demonstrator rank and the difference between demonstrator and participant rank (demonstrator rank: -0.35, [-0.48, -0.22]), rank difference: -0.06, [-0.17, 0.05], interaction: 0.036, [0.001, 0.026], **Fig. 3.2f**) - successful demonstrators were most likely to be copied, particularly when offered to a poorly performing participant. A participant's rank predicted their asocial performance (-0.13, [-0.21, -0.04]), showing that higher-ranking demonstrators offered higher-quality information. The minimum adequate model for each of the three tasks contained a baseline level of social information use and a random subject effect. The model for the “pitch task” also included the effect of demonstrator number and the cost to listen to the sounds. The “intensity task” minimal adequate model contained the effect of demonstrator

consensus whilst the “pitch modulation task” model also included demonstrator rank, the difference in rank between subject and demonstrator and the interaction between these two.

### 3.5 - Discussion

These experiments show that several factors simultaneously influence human social information use, consistent with a number of social learning strategies, and that such behaviour leads to adaptive decision making. The findings provide confirmatory support for all nine hypotheses derived from evolutionary models (Boyd and Richerson, 1985; Henrich and Boyd, 1998; Schlag, 1998; Boyd and Richerson, 1988).

#### 3.5.1 - The number of demonstrators

Three of four experiments found that the rate of copying increased with the number of demonstrators (**Fig. 3.2d**). As majority decisions of larger groups are known to be more likely to be correct than those of smaller groups (King and Cowlshaw, 2007), this tendency is likely to prove adaptive. From **figure 3.2d** it is clear that the impact of demonstrator number was enhanced in experiments that used the parallel protocol (i.e., when participants chose between social and asocial information, as opposed to receiving both in sequence). This difference cannot be attributed to a commitment

effect (Brody, 1965), where participants favour earlier decisions, as when the first experiment was repeated, but with the order of information presentation reversed, similar levels of social information use were observed. It is more likely that the weaker effect observed with the linear protocol is caused by the necessity to limit the analysis to cases where the participants' initial decisions disagreed with the majority of demonstrators - something that was not required when using the parallel protocol. It may also be the case that the trials on which the asocial and social information disagreed created a mistrust of the social information in participants, lowering their reliance on social information. As participants only received a single source of information on each trial under the parallel protocol such disagreement never occurred and so the distrust of social information would not have developed.

### 3.5.2 - Participant confidence

Subject confidence in their own judgement had a negative effect on the probability that they would use social information (**Fig. 3.2e**). This finding is consistent with experimental work using non-human animals, which reports a *copy-when-uncertain* rule in rats and nine-spined sticklebacks (van Bergen et al., 2004; Galef Jr. et al., 2008; Kendal et al., 2005). This behaviour was adaptive, in the sense that it increased participant pay-off, as participant confidence genuinely predicted whether participants were correct or not (**Fig. 3.2e**). Incorrect participants typically expressed lower confidence than correct participants, and so were more likely to use social information and change their answer to the correct choice.

### 3.5.3 - The cost of asocial information

I also found evidence consistent with a *copy-when-asocial-learning-is-costly* strategy (Boyd and Richerson, 1985). In the third experiment, participants were more likely to use social information when informed the upcoming trial was difficult than when it was an easy trial. Across trials, participants using asocial information received lower pay-offs on harder than they did on easier questions, and so this bias to rely on social information on more difficult trials allowed participants to avoid costly asocial information. Similarly, participants in the third experiment were also more likely to choose social information if they had incurred high costs using asocial information on the previous trial than if they had incurred a low cost. Finally, in the fourth experiment participants chose to use social information, instead of asocial information, more frequently when the cost to listen to the sounds was high compared to when it was low.

### 3.5.4 - Changes in social learning across trials

There was a drop in social information use across trial groups in Experiment 3, however, there was a simultaneous increase in asocial performance showing that this was an adaptive response. With practice, it appears participants justifiably became increasingly confident in their own abilities and subsequently decreased their reliance

on social information. I also observed an interaction with demonstrator number, such that social-information use decreased when there were 3 or less demonstrators, but remained steady when there were 5. This is consistent with participants changing their use of social information over time, potentially becoming aware of the risks of reliance on a small number of demonstrators (King and Cowlshaw, 2007), however, it can also be interpreted in the context of their own improving performance. It is possible that as participants become more confident in their own abilities they felt their own decisions offered more information than groups of three or fewer demonstrators, but that a group of 5 was still the better source of information.

### **3.5.5 - Different tasks**

No effect of task was found in the third experiment. This is in agreement with additional analyses directly comparing the results of the first two experiments that found no difference in the way that participants used social information. However, participant experience differed between the two experiments as participants in the second experiment typically expressed lower confidence (-0.16, [-0.27, -0.05]) and more frequently reached the wrong answer after having only received asocial information (-0.44, [-0.84, -0.05]) suggesting participants accurately perceived it to be a more difficult task. Thus any difference in the amount of social information used between the two tasks is due to the same decision making strategies being used on two tasks of different apparent difficulty, not the different natures of the tasks.

A difference was found, however, between the third and fourth experiments; only in the former did costs incurred when learning asocially have an impact on the likelihood that participants would subsequently use social information. Although this is hard to explain in light of the predictions of evolutionary models, explanations can be made based on methodological differences between the experiments. In the third experiment, participants were timed on how long they took to answer and incurred a cost proportional to this time that reduced their payoff. In the fourth experiment, however, participants incurred a cost every time they clicked to play a tone, the magnitude of which they were informed. As a result, in the third experiment, the cost was hard to anticipate without accurate information on how long participants had taken, whereas in the fourth participants were likely aware of how great a cost they had incurred. In the fourth experiment it is possible that participants justified their expenditure to themselves before being informed of it. They may also have viewed the costs as unavoidable, as in order to compare the two tones they would have to listen to them both at least once. In the third experiment, however, costs would have been unknown until participants were informed of them and so the surprise of large, unanticipated costs may have led to larger subsequent adjustments in behaviour.

### **3.5.6 - Payoff-biased copying**

In Experiment 4, participants' performance rank did not affect the likelihood of copying, but demonstrators' performance did (**Fig. 3.2f**). This is strikingly consistent with the pattern of pay-off-based copying observed in nine-spined sticklebacks (Pike

et al., 2010), which also copied in proportion to the success of the demonstrator, regardless of their own success. Thus humans and fish alike behave in accordance with “proportional observation” (Schlag, 1998). Selectively copying higher-performing demonstrators is likely to be adaptive, since such individuals are more likely to demonstrate behaviour associated with high returns (Laland, 2004; Boyd and Richerson, 1985). In our experiment participants' performance ranking did indeed predict their asocial performance, so top-ranking individuals were genuinely more skilled at the task and had not simply copied their way to the top. Participants were also sensitive to the demonstrator's performance relative to their own and were even more likely to copy when being outperformed than otherwise. However the presence of this effect was contingent on the demonstrator's ranking; it was most pronounced when the demonstrator was top ranking, but all participants were equally unlikely to copy poorly performing demonstrators, irrespective of their own performance. Thus information from strongly performing demonstrators was most attractive to poorly performing participants such that behaviour matched “proportional observation” with conditional “proportional imitation” (Schlag, 1998). It is interesting that the effect of relative performance is conditional on demonstrator performance as one might expect the decision of a bottom ranking individual to be more valuable to a second-from-bottom ranking individual than a top ranking individual. I suggest that the observed behaviour is the result of three additive motivations to copy: one to acquire high quality information, a second to acquire information of a better quality than you yourself could provide, and a third to avoid the cost of acquiring your own information. The first is what underlies the effect of demonstrator ranking - better ranking individuals offer better information and so are better prospects as demonstrators. The second explains the effect of relative performance - if an

individual outranks you then they offer better information than you are likely to acquire yourself, this effect disappears for bottom ranking demonstrators as their information is of a lower quality than everyone else's. Finally, the third motivation explains why even bottom ranking demonstrators have a non-zero probability of being copied - although they offer the poorest source of information, it is at least free and so represents a high-risk, high-reward option.

### **3.5.7 - Frequency dependent copying & conformist transmission**

Subjects were sensitive to the frequencies of the two options in the demonstrators' decisions (**Fig. 3.3**). The observed relationships produce a behavioural response closely matching the S-shaped curve of conformist transmission (Boyd and Richerson, 1985; Henrich and Boyd, 1998), but only when the number of demonstrators is large and participant confidence is very low (**Fig. 3.3a and b**). In other cases the S-shaped curve is shifted off to the right or flattened out. However, formal models of conformist transmission typically assume individuals choose between two variants without prior information, which was not the case under the linear protocol, as the asocial information constitutes prior information. Thus, a fairer test of whether humans show a conformist bias in their social learning is to examine their response to consensus in the absence of prior information. I achieved this through statistical controls, and in this case the effect of social information alone is clearly conformist, with a symmetric S-shaped curve (**Fig. 3.3c**). Thus, our data are consistent with a conformist response to social information, but as participants' behaviour is the result



of both social and asocial influences, the resultant behaviour may not match a conformist S-shaped curve, depending upon the magnitude of the asocial influence (**Fig. 3.3d**).

### **3.5.8 - A bias against social information**

The rarity of conditions under which conformist behaviour is realised, despite an underlying conformist response, suggests a bias towards asocial over social information. Increased participant confidence strengthens this bias, making participants more willing to go against the consensus (**Fig. 3.3a**). Such behavioural flexibility may explain the contradictory findings of previous work. For example, data from the Asch experiments found that participants copied at lower rates than predicted by conformist transmission models (Efferson et al., 2008). The simplicity of the task used likely leads to high participant confidence in asocial information, shifting the curve to a point where conformist behaviour would not be realised. Even against unanimous social information, participants may be unlikely to copy on tasks that they believe they can readily solve asocially and that even on difficult tasks, participants may not use social information if there are few demonstrators. Accordingly, I anticipate that conformist behaviour will depend on demonstrator number as well as the characteristics of the task.

A bias for asocial over social information has been documented elsewhere (van Bergen et al., 2004; Eriksson and Strimling, 2009; Weizsäcker, 2010) and various explanations have been proposed for it. These include that (1) people have a distrust of socially acquired information, (2) asocially acquired information has a greater learning impact, and (3) people have prosocial tendencies that lead them to produce information (Eriksson and Strimling, 2009). I offer three further suggestions; that (4) the tasks thus far deployed have been sufficiently simple that participants have considerable confidence in their own judgement, (5) the low cost of making an incorrect answer favours asocial information as, although potentially the poorer source of information, participants found it more enjoyable (several individuals from our experiment reported finding “having-a-go” rewarding), and (6) that such a bias is the product of a mixed population of ‘conformists’ (who readily use social information) and ‘mavericks’ (who resist social information) (Efferson et al., 2008; Mesoudi, 2011). With regards to the latter, this should not be interpreted as two types of individual, but instead as a continuum of reliance on social information across individuals, ranging from heavily reliant ‘conformists’ to much less reliant ‘mavericks’. In agreement with this, our analysis finds participants rely on social information to different extents (precision of population distributions, experiment 1: 0.27, [0.10, 0.73]; experiment 2: 0.26, [0.13, 0.51]; experiment 3: 0.44, [0.24, 0.78]; experiment 4: pitch task: 0.21, [0.12, 0.35]; intensity task: 0.16, [0.10, 0.25]; modulation task: 0.15, [0.10, 0.26]), but no evidence of discrete personality types.

### 3.5.9 - Conclusions

In summary, these experiments provide conditional or strong support for the deployment of several social learning strategies predicted by the theoretical cultural evolution literature, including conformist transmission, pay-off based copying, *copy-when-asocial-learning-is-costly* and *copy-when-uncertain* (Laland, 2004; Boyd and Richerson, 1985; Henrich and Boyd, 1998; Schlag, 1998, 1999; Kendal et al., 2009). Importantly, these various influences appear to operate simultaneously, and interact to produce behaviour leading to effective decision making and higher pay-offs. These results are not incompatible with the notion that social learning may sometimes have maladaptive consequences (Tanaka et al., 2009), but nonetheless illustrate the adaptive aspects of cultural transmission necessary for it to have been favoured by selection. It is apparent that formal evolutionary models provide a framework for predicting the conditions that are likely to produce a reliance on social information on which maladaptive cultural traits can occasionally ride (Boyd and Richerson, 1985). These results support the effectiveness of the cultural evolution theory in predicting human copying behaviour.

## Chapter 4

# **An Empirical Study of Social Learning Strategies in Young Children**

## **Abstract**

Human culture relies on extensive use of social transmission, which, as shown in Chapters 1 and 2, must be integrated with asocially acquired knowledge for effective decision-making. Formal evolutionary theory predicts that natural selection should favour adaptive learning strategies, including a bias to copy when uncertain, and a bias to disproportionately copy the majority (known as conformist transmission, see Chapter 2). While the function and causation of these evolved strategies is comparatively well-studied (see Chapter 3), less is known of their development. In this chapter, I present an experimental investigation into the development of the bias to copy-when-uncertain and conformist transmission in children from the ages of 3 to 7, testing predictions derived from theoretical models. Children first attempted to solve a binary-choice task themselves, but were then given the decisions of informants and an opportunity to revise their answer. I investigated whether the children's reliance on social information was contingent on (i) the quality of asocial information and (ii) the consensus amongst informants. As predicted older but not younger children copied others more when uncertain, though they also showed a tendency to stick with their initial decision. I also found that older children, like adults, were disproportionately receptive to non-total majorities (i.e., were conformist) whereas younger children were receptive only to total majorities (i.e., were anti-conformist). These findings imply that adult-like adaptive social information use develops between the ages of 3 and 7.

## **4.1 - Introduction**

Given the selective trust by young children detailed in **Chapter 1**, it is clear that a wide variety of factors can influence whether or not a child relies on the information offered by another individual, and that such trust changes across childhood, particularly between the ages of 3 and 4. Although the findings of developmental psychology are of great value to cultural evolutionists, the fields do not share a common theoretical framework guiding experiments and several of the learning strategies investigated in adults are unstudied in children. The work presented here seeks to use the well-established experimental methods of developmental psychology to investigate the presence of a bias to copy others when uncertain and conformist transmission in young children.

### **4.1.1 – Copy-when-uncertain**

A bias to copy when uncertain has been a key assumption of theoretical models of cultural evolution (Boyd and Richerson, 1985; Enquist et al., 2007). It has also been documented in adult humans (see **Chapter 3**) as well as a range of non-human taxa (van Bergen et al., 2004; Galef Jr. et al., 2008). Despite this, there are no investigations of the development of an adaptive learning mechanism guided by uncertainty. However, there are several findings that provide indirect evidence that such a bias is present in young children and it is well-established that infants and young children have at least some sensitivity to confidence and use it to guide their social learning. For example, twelve- and 16-month-olds look more rapidly and more often at nearby adults and are more receptive to affective signals from the adults,

when presented with an ambiguous (e.g., unfamiliar or strange), as opposed to an unambiguous, toy (Harris and Lane, in press). Their emotional reactions toward the ambiguous toy are likely to reflect the signals they receive – positive in the context of happy signals but negative in the context of fearful signals. They also approach or withdraw from the ambiguous toy depending on such signals. Such adjustments of emotion and behaviour do not occur, however, if the toy is unambiguously positive or negative (Kim and Kwak, 2011), showing that the socially guided behaviour is triggered by the uncertainty of the situation. A similar pattern is seen when 18-month-olds face an intermediate slope. Whether they walk down the slope or remain immobile depends on whether their mother's affective signals are positive or negative. Yet if the slope is unambiguously gentle or steep, maternal input has little impact; children advance down gentler slopes but baulk at steeper slopes irrespective of maternal signals (Tamis-LeMonda et al., 2008).

#### **4.1.2 – Conformist transmission**

Conformist transmission (as detailed in **Chapter 2**) is the disproportionately large influence of majorities on an individual's decision making. Theoretical models suggest conformist transmission is a highly effective strategy across a range of environmental conditions (Boyd and Richerson, 1985; Nakahashi et al., 2012; Eriksson et al., 2007) and in **Chapter 3** I presented empirical data showing conformist transmission in the decision making behaviour of adults. Again, although there is currently no data on the development of conformist transmission, there is evidence consistent with such a strategy at an early age. For example, studies with preschool children have shown that they are more likely to rely on information from a group of

adults in apparent consensus than from a single adult. When presented with conflicting names for a novel object by two different informants, 4-year-olds are receptive to the reactions of bystanders. If two bystanders signal agreement (via head nods and smiles) with one name and disagreement (via head shakes and frowns) with the other name, children overwhelmingly endorse the name eliciting bystander agreement (Fusaro and Harris, 2008). Similarly, if three informants all point to the same object as the referent of a novel name whereas a single informant points to a different object, 3- and 4-year-olds select the former when asked to identify the named object (Corriveau et al., 2009a). Three and 4-year-olds also defer to a consensus of three informants who produce judgments that conflict with what children can see for themselves. Having correctly identified the biggest line of a trio, children will change their judgment following a different claim by the consensus (Corriveau and Harris, 2010) at rates similar to those observed in classic studies of conformity in adults (Asch, 1956; Bond and Smith, 1996). Thus, children often endorse a consensus when they lack relevant perceptual cues (as in learning names for novel objects) but they will even do so despite the availability of perceptual cues.

#### **4.1.3 – Outstanding questions**

Taken together, these findings show that young children are more likely to seek out and accept guidance from an adult when the available information is ambiguous. In addition, they favour guidance from a consensus rather than a single informant. Nevertheless, this body of research displays three limitations. First, the quality of the information available to children has not been systematically varied. We do not know if children's reliance on social information is sensitive to small changes in quality or if



they use social information only when asocial information is below a certain quality. Second, the nature of the consensus has not been systematically varied. In particular, children might or might not be sensitive to the size of the majority relative to the minority. Thirdly, we do not know the relationship between these two social learning biases. Each might guide children's social learning independently or they might interact with one another. Finally, the adaptive value of childhood social learning has not been considered.

There are three reasons to anticipate that developmental changes in social learning may occur (Harris, 2012). Firstly, strategies that are adaptive for adults may not be for children, in which cases adaptive age-specific social learning strategies may have been favoured by selection. For example, given their naivety and vulnerability, it may be adaptive for young children to acquire knowledge from particular sources (e.g. parents, kin, familiar individuals) who are likely to be reliable, whilst older, more experienced and less vulnerable individuals may benefit from attending to alternative sources (e.g. experts) (Boyd and Richerson, 1985; Henrich and Henrich, 2010). Secondly, if the development of social learning strategies is contingent on experience of social groups, we should expect strategy usage to change as a child's social horizon expands. As such experiences may differ between populations, such a mechanism could result in cultural differences in the development of social learning. Thirdly, very young children may be incapable of some forms of computation (e.g. computing the majority behaviour) necessary to implement adaptive social learning strategies (e.g. conformist transmission), rendering their independent decision making suboptimal prior to a critical developmental stage. The three cases described above offer different (although not mutually exclusive) explanations for changes in behaviour across

development; in the first case alternative strategies have been selected, in the second they emerge depending on relevant experience, whilst in the third case they are explained by the presence/absence of other cognitive processes. On the basis of the aforementioned evolutionary models, I predict that any changes observed in children's behaviour as they age should be leading towards the behaviour seen in adults. Additionally, I predict that such changes should lead to an increase in the adaptive value of such behaviour. Although such a prediction may seem highly likely, the presence of adaptations specific to childhood could result in the behaviour of children initially diverging from that of adults, before later converging.

To evaluate these predictions and questions, I present an experimental study in which children were given a task that they first attempted to solve themselves, but were then informed of the decisions of a group of adults and given the opportunity to revise their decision. I investigated the effect of individual confidence (inferred from manipulations of task difficulty), as well as the nature (as opposed to merely the presence) of an effect of informant consensus on children's decision making.

## **4.2 - Methods**

### **4.2.1 - General Methods**

I carried out a single experiment in which young children took part in a computer-based, binary choice game using asocial and social information to make relative quantity judgements. Children gained asocial information through direct access to the task, whilst the social information gave the decisions of 10 adult informants. Each

participating child completed 5 trials, taking 5 minutes, and was rewarded with a sticker for taking part irrespective of their performance.

#### **4.2.2 - Subjects & Apparatus**

122 children took part, 55 were male and 67 female. The children were between 2 years 11 months and 8 years 11 months (mean age = 5 years 7 months, median = 5 years 5 months). The experiment took part in the “Discovery Centre” in the Museum of Science, Boston, and children were recruited from visiting families. Children played the game individually, although a parent/guardian was present throughout.

#### **4.2.3 - The Task**

The experiment used the “who-has-more” two-alternative forced-choice (2AFC) numerical discrimination task in which the child must decide which of two television characters (Big Bird or Grover) has more dots. These characters were chosen because they are typically familiar to children, if a child did not recognise the characters the experimenter “introduced” them. This task was used because previous work has established how to manipulate the difficulty of the task for children, with difficulty related to the similarity in the number of dots each character has (Halberda and Feigenson, 2008). This can be expressed as a dot ratio, calculated as the difference between the numbers of dots each character has, divided by the lesser. As the dots ratio tends to 0, the trial becomes increasingly difficult. In such perceptual tasks, the confidence associated with decisions is known to be robustly related to difficulty, with decisions on more difficult trials made with lower confidence (Pleskac and

Busemeyer, 2010). Furthermore, children are known to be sensitive to the difficulty of each trial (Odic et al., 2012). Thus, I inferred that the confidence of participating children would be associated with task difficulty, and so by manipulating difficulty, I could influence confidence.

On each trial, the dot ratio was between 0 and 1 and each character had a random number of dots between 10 and 30 (though there was always at least 1 dot difference between the two characters and all ratios between 0 and 1 were equally likely). The minimum of 10 was used as for these numbers the dot ratio is all that is needed to predict difficulty (given the age of the child), whereas for lower numbers (<5) individuals use different enumeration mechanisms for which ratio does not predict difficulty. The location of each dot on its panel was randomised, and no dots overlapped. The dots were shown for 3.5s as trials with a dot ratio close to 0 were still very hard after 3.5s, whilst trials with a great dot ratio were relatively simple. To avoid the total area of a character's dots serving as a cue to the total number of dots, as I wanted children to respond to number alone, dots were resized using an area anti-correlation procedure to prevent total area being a reliable cue (Halberda and Feigenson, 2008). Using this procedure, each trial had a 0.5 chance of being anti area-correlated in which case the relationship between the number of dots and the total area covered was reversed such that if one character had twice as many dots as the other character, the sum of their dots' areas was half that of the other character's dots. Additionally, the diameter of each individual dot was multiplied by a number drawn from a uniform distribution ranging from 0.65 to 1.35 to add variation in size.

#### **4.2.4 - The Social Information**

The social information was presented as a video of 10 informants all of whom thought one of the characters had more dots than the other. During each video, a voice-over asked the informants if they thought each character had more dots (e.g., “Who thinks Grover has more?”). At each asking, some informants in the video nodded (a signal children are known to recognise, Fusaro and Harris, 2013) and raised their right hand to signal agreement whilst the others looked down and remained still to signal disagreement. The informants were in two rows of 5 (the front row was seated) and all were women wearing identical purple t-shirts and without any identifying items (e.g., glasses). I made 4 repeats of each possible video (with 0 to 10 of the ten informants supporting each option) with the spatial arrangement of informants varying across videos such that each informant did not occupy a consistent location. The intention was that children playing the game would not be able to recognise any informants across trials to prevent them from trusting specific individuals.

#### **4.2.5 - Procedure**

Children joined the experimenter in the experimental area of the “Discovery Centre”. The experimenter explained to the child how to play the game and then guided them through it without leading their decision making. For each trial the child was first shown two side-by-side, grey rectangular panels in which the dots were to appear. An image of each character was in front of their panel (which panel was Grover's and which was Big Bird's was randomised for each trial) and the panel was outlined with a border of "their" colour (blue for Grover, yellow for Big Bird). The dots (also in their

character's colour) then appeared for 3.5s. After the dots had disappeared the child was asked who they thought had more dots. After the child made their initial “asocial” decision, a video of 10 informants was played to provide social information. The video shown on each trial was randomly selected with respect to both the child's initial decision and the trial dot ratio. The voice-over always asked the informants about the character on the left first. After the video, children were again asked to make a decision and the trial was complete. To avoid confidence hysteresis, children were not given feedback on their performance during the experiment . However, after all 5 trials a final screen congratulated the child, informing them they had done "really well" (irrespective of the child's actual performance) and the experiment finished. When the game was complete the child was given a sticker and they were free to go.

### 4.3 - Analysis

The analysis consisted of two parts. I analysed the data using generalized linear mixed models with a Bernoulli error structure and logit link function, using Markov Chain Monte Carlo methods to fit the models and generate credible intervals in OpenBUGS 3.2.1 (Lunn and Spiegelhalter, 2009). All priors used were extremely uninformative, for more details see **section 4.3.1**. Minimal adequate models were constructed by backwards elimination, removing variables for which the **90%** credible interval included 0. Parameter values were estimated using a sample of at least 3000 iterations, after a suitable burn-in period and thinning to remove autocorrelations. Unless otherwise stated, all graphs show the median and **95%** credible interval of the posterior sample from the fitted models as estimates of parameter values. For a comparison of model estimates and raw data averages, see **Figure 4.1**.

The first part of the analysis modelled the probability that a child's initial decision (following asocial information only) would be correct ( $p_{IDC}$ ) with additive effects of dot ratio ( $DR$ ), an independent performance effect ( $\beta_2$ ) and a categorical effect of which side of the screen the character with the most dots was displayed on, such that:

$$f(p_{IDC}) = \beta_1 * DR + \beta_2 + \beta_3 * side\ of\ screen, \quad (4.1)$$

where  $f$  is the logit link function and  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  are coefficients. The independent performance effect ( $\beta_2$ ) should only have a non-zero value if  $DR$  is unable to account for subject performance, whilst the screen side effect allowed children to favour one side of the screen over the other.  $f(p_{IDC})$  can be thought of, and will henceforth be referred to, as the “net evidence” children were able to collect.  $DR$  contained a linear effect of dot ratio that interacted with age (linear), sex (categorical), whether the trial was area-correlated or not (categorical) and random individual effects, such that:

$$DR = dot\ ratio * (\beta_4 * age + \beta_5 * sex + \beta_6 * area\ correlation + individual\ effects), \quad (4.2)$$

where  $\beta_4$ ,  $\beta_5$ , and  $\beta_6$  are coefficients, and a dot ratio of 0 (hypothetical case where both characters have the same number of dots) causes  $DR$  to have a value of 0, meaning that children gained no net evidence. The interactions allowed performance across trials to change with age, sex, area correlation and individual. As part of the backwards elimination procedure, the following parameters were removed from the final model: the screen side bias, independent performance effect, the interaction

between area correlation and dot ratio and the interaction between sex and dot ratio (i.e.,  $\beta_2=\beta_3=\beta_5=\beta_6=0$ ).

The second part of the analysis modelled the probability that a child's final decision, after asocial and social information, would be correct ( $p_{FDC}$ ) with additive effects of the net evidence collected asocially ( $f(p_{IDC})$ ), a tendency to stick with your initial decision ( $IDC$ ) and the social information ( $SI$ ) such that:

$$f(p_{FDC}) = (\beta_7 + \beta_8 * age) * f(p_{IDC}) + (\beta_9 + \beta_{10} * age) * IDC + (\beta_{11} + \beta_{12} * age + \beta_{13} * sex + \beta_{14} * dot\ ratio + individual\ effects) * SI, \quad (4.3)$$

where  $f$  is the logit link function and  $\beta_{7:14}$  are coefficients. The value of  $f(p_{IDC})$  is the same as in equation (4.1). This value not only contains information about the decision children initially made, but also the net evidence they based it on. If  $(\beta_7 + \beta_8 * age) = 1$ , that would represent children bringing the full amount of their asocial information forward into the second decision. Conversely, if  $(\beta_7 + \beta_8 * age) < 1$  then only part of the information is brought forward (perhaps as the result of forgetting) and if  $(\beta_7 + \beta_8 * age) > 1$  the prior net evidence is inflated. (Note that negative values, although permitted in the model, have no obvious interpretation). The sticking tendency allowed individuals to be more (or less) likely to stick with their initial decision than the net evidence they collected would predict. The effect of the social information was calculated such that:

$$SI = q^s / (q^s + (1-q)^s), \quad (4.4)$$



where  $q$  is the proportion of informants who are correct and  $s$  is the shape parameter such that:

$$s = \exp(\beta_{15} + \beta_{16} * \text{age}), \quad (4.5)$$

where  $\beta_{15}$ , and  $\beta_{16}$  are coefficients.  $SI$  is monotonic and has a value of 0 when there was no consensus ( $q=0.5$ ), and a maximal magnitude when all 10 informants chose a particular option ( $q=0$  or 1). According to the value of the shape parameter,  $s$ , which interacted with age, the function could produce an enlarged or diminished response to non-total majorities. The following parameters were removed from the final model: the interaction between age and  $DR$ , the interactions between sex, dot ratio and age and  $SI$  (i.e.,  $\beta_8=\beta_{12}=\beta_{13}=\beta_{14}=0$ ).

#### 4.3.1 - Priors

Priors for all parameters described above are detailed below. Normal distributions are specified as  $N(\text{mean}, \text{precision})$  and gamma distributions are specified as  $G(\text{shape}, \text{rate})$ , where precision is  $1/\text{variance}$  and rate is  $1/\text{scale}$ .

$$\beta_{1:14} = N_{1:14}(0.0, 0.01)$$

$$\beta_{15,16} = N_{15,16}(0.0, 1)$$

individual effect  $_{e,i} = N_{e,i}[0, G_e(0.01, 0.01)]$ , for individual  $i$  and effect  $e$  (i.e., whether the individual effect was on asocial performance, or social information use) and where  $G_e()$  is the prior for the precision of the random effect.

The priors for  $\beta_{15,16}$  are more informative than for the other parameters as these values were combined and then exponentiated (see equation 4.5) such that even these comparatively precise values allowed extreme results and any less precise priors lead to the exponential producing infinitely large values, despite variables being standardised.

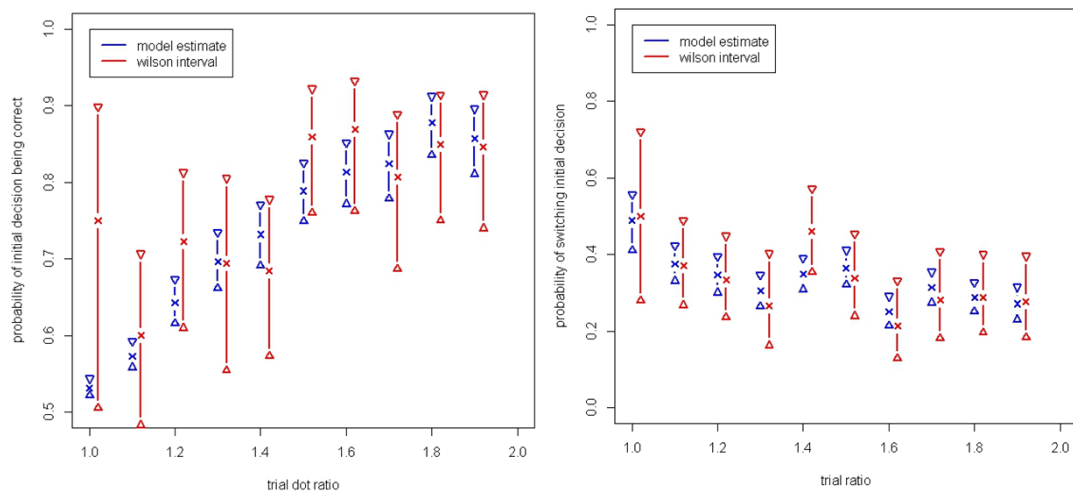


Figure 4.1

These figures compare the model estimates of the mean of the raw data (in blue) with the actual mean of the raw data (in red). In both panels the red lines are the actual mean and Wilson binomial confidence interval (a simple method of generating confidence intervals for binomial data) whilst the blue lines are the median model estimate of the mean of the raw data and its 95% central credible interval. The similarity between the red and blue intervals across the two figures shows that our model was well able to fit the data. Meanwhile, the observation that the model intervals are narrower than the Wilson intervals highlights the increase in the precision of estimates facilitate by the analyses.

## 4.4 - Results

Unless otherwise stated, all quoted effects are medians, followed by the corresponding 95% central credible intervals.

Following asocial information only, children performed much better on trials with a high (i.e. easier) rather than low (i.e. harder) dot ratio ( $\beta_1$ , 3.16, [2.53, 3.95]) and the

magnitude of this improvement increased with age ( $\beta_4$ , 0.89, [0.50, 1.38], see **Fig. 4.2a**). There was no evidence of a difference in performance between girls and boys ( $\beta_5$ , -0.15, [-0.13, 0.96]) and no evidence that area correlation helped performance ( $\beta_6$ , 0.75, [-0.19, 1.71]). There was also no performance effect independent of dot ratio ( $\beta_2$ , 0.25, [-0.13, 0.62]) suggesting the effect of dot ratio and its interactions were able to account for performance. Children did not, as a group, show a side preference ( $\beta_3$ , -0.13, [-0.34, 0.08]).

When given social information, children showed a tendency to “stick” with their initial decision ( $\beta_9$ , 1.96, [1.30, 2.61]), regardless of whether they had been right or wrong, and independent of the net evidence they had collected. This tendency to stick increased with age ( $\beta_{10}$ , 0.68, [0.25, 1.12], see **Fig. 4.2b**).

However, in addition to the above, we also find strong evidence that children took the magnitude of their own net evidence into account ( $\beta_7$ , 0.29, [0.11, 0.50], see positive gradient of lines in **Fig. 4.2b**), although it was undervalued relative to what their individual performance would suggest. Moreover, there was no strong evidence that this sensitivity to their own net evidence changed with age ( $\beta_8$ , 0.10, [-0.03, 0.24]). These two effects are independent and additive, such that although collecting more net evidence makes it more likely that a child will stick, even with very little net evidence children will tend to stick with their initial decision.

Finally, children were clearly influenced by the information provided by the informants ( $\beta_{11}$ , 0.32, [0.21, 0.53], see **Fig. 4.2c**). Their response to total majorities did not change with age ( $\beta_{12}$ , 0.05, [-0.02, 0.05]), sex ( $\beta_{13}$ , 0.05, [-0.17, 0.29]) or dot

ratio ( $\beta_{14}$ , 0.11, [-0.17, 0.40]). However, their response to varying levels of consensus ( $\beta_{15}$ , -0.84, [-1.49, 0.14]) did change with age ( $\beta_{16}$ , 1.00, [0.62, 1.50]) such that younger children were anti-conformist, but older children were conformist (see **Fig. 4.2d**).

The younger children effectively treat the social information as having three configurations; (i) complete endorsement (10 vs. 0), in which case they stick, (ii) complete opposition (0 vs. 10), in which case they change their decision or (iii) disagreement, in which case they show little to no adjustment in their behaviour regardless of the relative sizes of the majority and minority. However, for the 7-year-olds, non-total majorities have a relatively large influence. In this case, increasing the size of the majority further strengthens its influence, but with diminishing returns.

There was considerable individual variation in asocial performance (precision of population distribution: 0.33, [0.14, 1.32]), but much less variation in the response to social information (precision of population distribution: 11.0, [3.37, 31.8]).

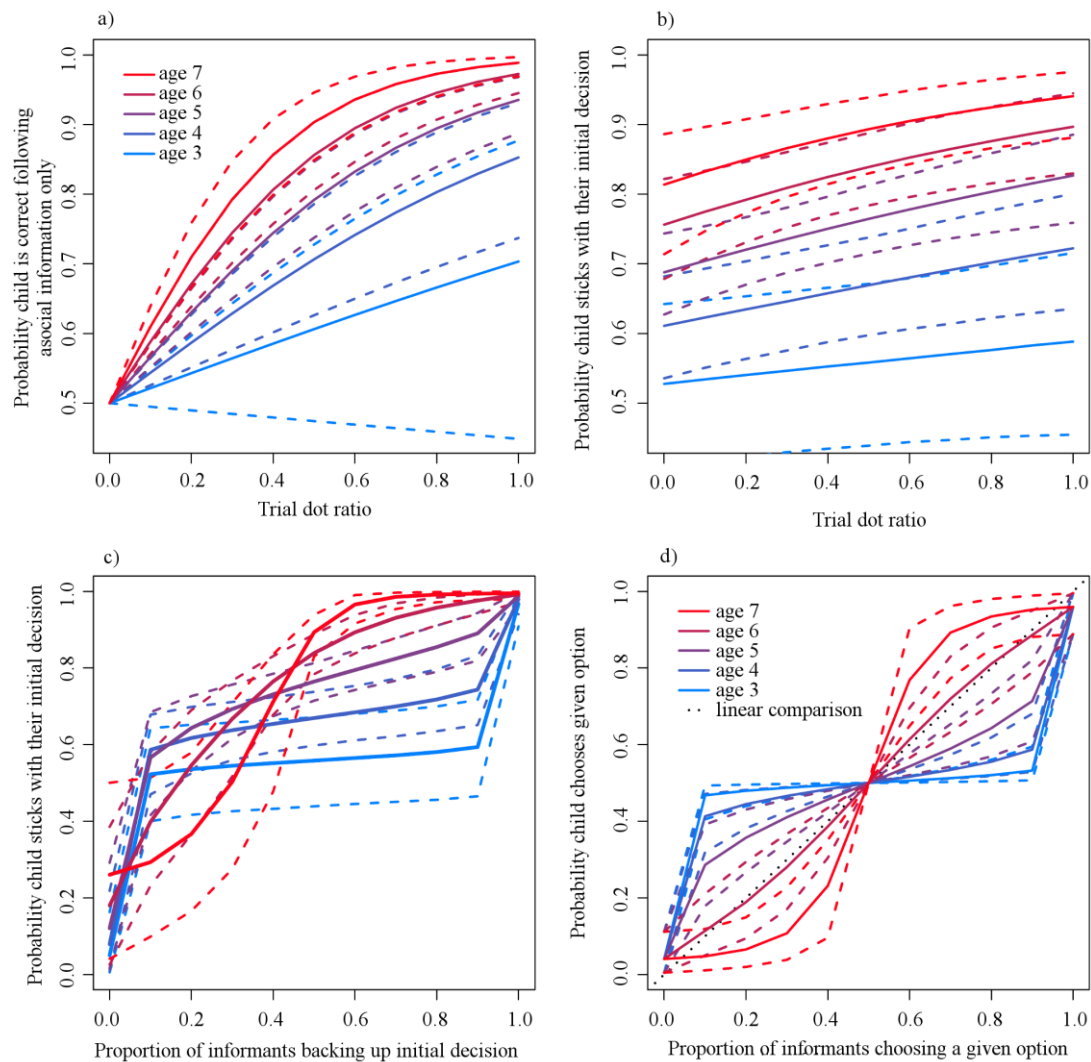


Figure 4.2

Figures show median estimates (solid lines), and 95% central credible intervals (dashed lines). (a) Children's performance improved with dot ratio and with age. Older children start to hit ceiling performance at intermediate dot ratios, whilst there is not strong evidence that the youngest children perform above chance levels. (b) The probability that a child sticks with their initial decision as a function of their initial net evidence and their sticking tendency (i.e., not the additional effect of the social information). Children showed a blunt tendency to stick with their initial decision at all dot ratios and hence irrespective of how much net evidence they had collected. This tendency to stick increased with age; the oldest children always have a >80% chance of sticking, whilst the behaviour of the youngest children is consistent with deciding at random. Nevertheless, children did show some sensitivity to how much net evidence they had collected themselves, being more likely to stick on trials with a high than low dot ratio. (c) The probability that children stick with their initial decision for a trial with the intermediate dot ratio of 0.5. Younger children, especially 3- and 4-year-olds are only affected by social information when there is unanimity amongst informants. However, older children, especially 6- and 7-year-olds, show a more nuanced response to social information and can discriminate between different non-total majorities. (d) The response of children to the social information alone (controlling for individual net evidence statistically). The black dotted line has a gradient of 1 (representing unbiased copying) and is for comparison with the other lines. Three-, 4- and 5-year-olds are anti-conformist in that they are at least somewhat insensitive to non-total majorities. Six-year-olds show a roughly proportionate response to the

size of the majority. Seven-year-olds, by contrast, are conformist in that they show an over-proportionate sensitivity to small majorities.

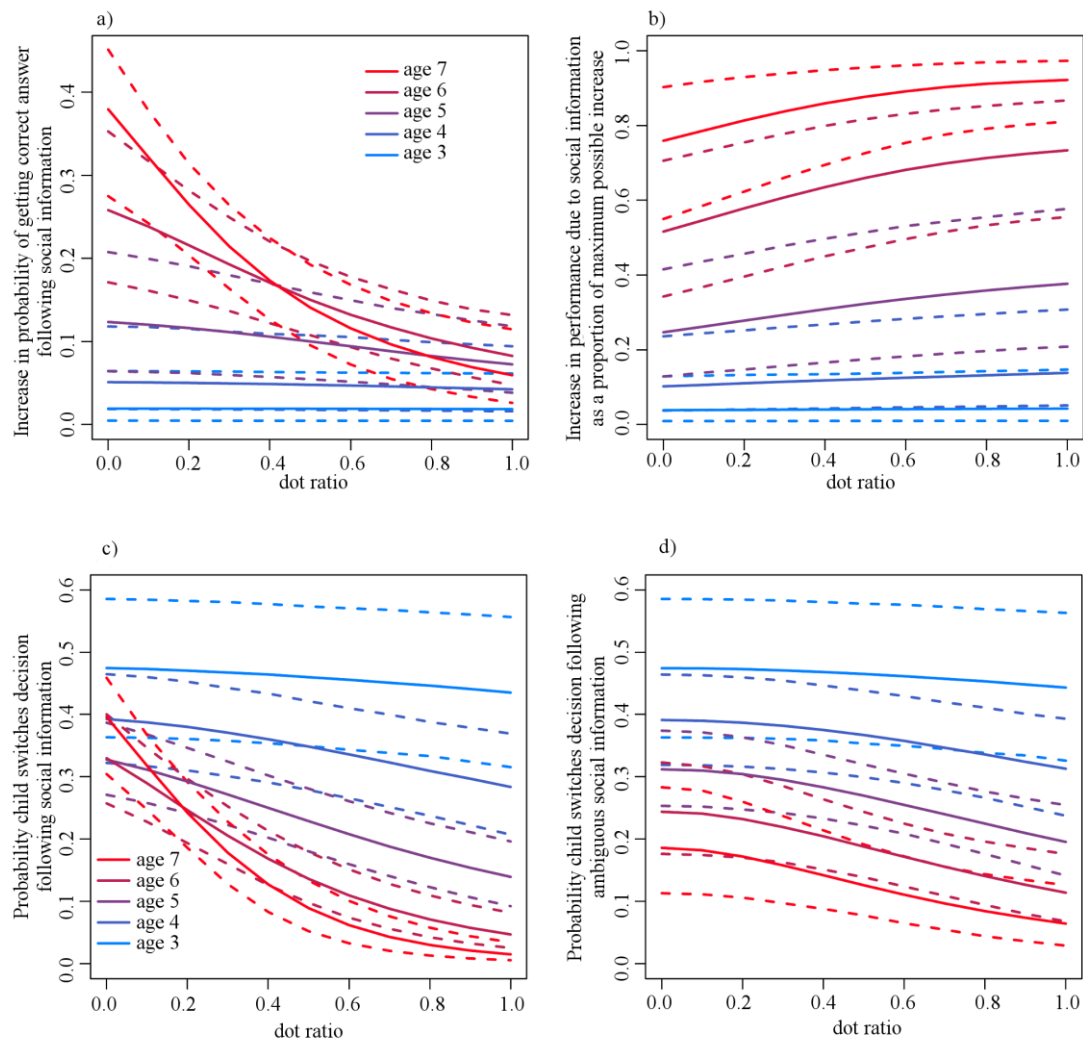


Figure 4.3

Figures show median estimates (solid lines), and 95% central credible intervals (dashed lines). (a) Given that 8 out of the 10 informants give the correct answer, older children were able to take advantage of the social information and used it to increase their performance, particularly on the more difficult trials. For easier trials the increase in performance due to social information was similar across ages, however, this is because on such trials older children are close to ceiling performance and so there is little room for further improvement. (b) In support of this, older children nearly maximised their performance following social information, particularly on easier trials, whilst the youngest children take minimal advantage of the social information at all ratios. (c) Older children, but not younger children, were more likely to switch following conflicting social information more difficult trials. The youngest children always had a high likelihood of switching, consistent with random behaviour. The graph is for the case of 8 out of the 10 informants disagreeing with the child. (d) Following social information without a majority (i.e., 5v5 informants), older children are still more likely to change their decision on more difficult trials. This tendency is much smaller than when the informants disagree with the children (panel c) and is likely due children doubting their decisions on harder questions. The extent to which this sensitivity to difficulty dictates switching matches the extent to which difficulty affects asocial performance (Fig. 4.2a).

## 4.5 - Discussion

### 4.5.1 - Consensus and conformist transmission

Children took part in a relative quantity estimation task in which they made an initial decision after direct access to the two quantities, but were then shown the decisions of ten adult informants and given the opportunity to revise their decision. My main finding is that the social information had a large effect on children of all ages. Indeed, the maximum magnitude of this effect did not vary with age as shown by the absence of any interaction between age and social information. The analysis suggests that in the absence of prior information, a child exposed to a total majority (i.e., 10 v 0) has a 90% chance of endorsing that judgement. However, the shape of the response to the consensus amongst informants changed sharply with age. The youngest children (ages 3-4) showed strong anti-conformism, where, although a total majority had a strong effect, all non-total majorities had little to no effect. By contrast, the older children (age 7) were conformist, showing an enlarged response to non-total majorities (although not as strong a response as to a total majority). The conformism of older children is an adaptive decision making mechanism predicted by the theoretical literature (Boyd and Richerson, 1985; Nakahashi et al., 2012; Kandler and Laland, 2013). It is also the same response to consensus I observed in adults in **Chapter 3**. However, the anti-conformism of younger children is harder to explain because it is not clearly of benefit to the children. I suggest that rather than it being a deliberate or selected strategy for young children to be anti-conformist, it is actually the result of their inability to respond appropriately to majorities that are less than unanimous. I do not claim that children are unable to recognise the presence of these majorities;

existing data shows even 3-year-olds should have no problem discerning the levels of consensus I used (Halberda and Feigenson, 2008). Instead, I suggest young children are unable to respond appropriately to a non-total majority even though they can recognise its existence. Whether young children see a group of 5 vs. 5 informants or 1 vs. 9 informants, although they can recognise that the consensus differs between groups, they register only that there is disagreement and leave their judgement unaltered. Regardless of the cause of this insensitivity to small majorities, these findings bear out my prediction that any changes observed across childhood should be towards increasingly adaptive and adult-like behaviour.

This result fits well with existing research finding other social behaviour develops across the same period. For example, 3 and 4-year-olds do not discriminate between two choices with identical rewards to themselves, but different payoffs to a partner. However, above the age of 5 children do discriminate and also show contingent reciprocity in that they reward partners who previously behaved cooperatively, but punish those who did not (House et al., 2013). In another study, whilst children between the ages of 3 and 8 endorsed norms for sharing, only 7- and 8-year-olds actually shared when the opportunity arose. Younger children even predicted that they wouldn't share when given the opportunity, ruling out the possibility that the lack of sharing was due to a failure of willpower (Smith et al., 2013). Together, these results illustrate a transition between the ages of 3 and 8, from individual focussed behaviour where sharing norms, payoffs to other individuals and intermediate majorities are recognised, but ignored, to increasingly socially guided behaviour where more weight is given to extra-individual factors. These observations concerning childhood cognitive development are not necessarily underpinned by the same cognitive



mechanisms, but nonetheless have qualitative similarities, appear on similar developmental trajectories and may be influenced by similar experiential factors.

#### **4.5.2 - Confidence and information quality**

I also varied the quality of information available to children in order to manipulate their confidence – a variable predicted by theory to influence social learning and observe to do so in adults as well as non-human species. First, I find that that children show a somewhat diminished sensitivity to the magnitude of their net evidence (the median estimate is that only 25% of the information was carried forward, see **Fig. 4.2b**) and there was only very weak evidence that this ability increased with age. It is possible that children are forgetting some of their net evidence, perhaps the result of the time gap between the two decisions (~15s) or the cognitive demands of having to combine delayed social information with prior asocial input. If this is the case, similar experiments with a different temporal structure might not find these effects. However, the lack of any improvement in this regard with age is problematic for such an explanation because I would expect children's memory to improve with age. Furthermore, in addition to, and independent of, the net evidence brought forward, children had a tendency to “stick” with their initial decision, regardless of how much net evidence they had to back it up, which did increase with age. For the youngest children this tendency was sufficiently weak as to be negligible. However, it was quite powerful among the older children (see **Fig. 4.2b**). Rather than forgetting their earlier decision, children seem to collapse their information down into a single binary choice which they increasingly weight with age, largely forgetting the magnitude of the net evidence they have collected. For easier trials (ratio > 0.5), this is not too far from

optimal because the “sticking” bias approximates the likely quality of your own information (compare **Figs. 4.2a and b**). On lower dot ratio trials, however, the “sticking” bias leads to an over-confidence in the child’s initial decision that increases with age. This is surprising because it suggests performance becomes less optimal with age, contradicting my hypothesis. However, observation of model estimates of performance across trials of different difficulty clearly shows that despite the increased sticking tendency, the adaptive value of social learning increases across childhood (see **Figs. 4.3a and b**), with the sticking tendency of older children being overcome by their increased sensitivity to non-total majorities. Additionally, the observed sticking tendency is similar to adult behaviour, where an individual’s own decision is consistently weighted as more valuable than the decisions of others (Mesoudi, 2011; Yaniv, 2004; Bonaccio and Dalal, 2006; Weizsäcker, 2010).

A difference between adults and children may lie in the strength of the bias to copy when uncertain. As I show in **Chapter 3**, although adults do tend to stick, their confidence has a large effect on their social information use, with very uncertain individuals likely to follow group information. However, in this study, the quality of a child’s information has a comparatively muted effect. Although it is certainly possible that the development of social learning may not be complete over the ages that we studied, it is worth noting that I am inferring the confidence of the children from their asocial performance, not asking for them to rate it directly as was done with adults. Inferring confidence from performance in this fashion might seem like a reasonable step. However, in adult studies where data on both confidence and performance have been collected the correlation between the two is far from perfect (Luna and Martín-Luengo, 2012). Thus, the relatively weak effect of the magnitude of net evidence

collected on children's second decision could reflect a generally poor ability on the part of humans to accurately assess their own performance.

A possible criticism of our design is that, as children always heard from the informants, we cannot differentiate between children changing their mind due to social influence or due to doubt in, or reflection on, their initial decision. However, by looking at cases where there was no majority amongst the informants (i.e., 5v5) we can observe the rate of switching independent of a social consensus (**Fig. 4.3d**), which would include factors such as doubt. The rate of switching in this case is considerably lower than when a majority of informants disagree with the child (**Fig. 4.3c**; 8v2) so we can be confident that the disagreement influenced behaviour.

#### **4.5.3 - The adaptive value of children's social learning**

A central finding of this study is thus the increase in the adaptive value of social learning across childhood. I found that, when a majority of informants disagree with them, young children are no more or less likely to switch their initial decision according to the difficulty of the trial. Older children, however, do show this sensitivity (see **Fig. 4.3c**). This is clearly adaptive, as trial difficulty strongly predicted asocial performance, so a bias to switch on difficult trials when others disagree with you will make you disproportionately likely to switch away from an incorrect answer. I can be confident that this switching is due to the social information and not other factors such as doubt in the initial decision as the rate of switching, amongst older children, is much lower when there is no majority in the social information (i.e., 5v5 informants, see **Fig. 4.3d**). Even in this case, however, the switching of older children

is contingent on difficulty and more likely on harder questions than easier questions. This effect is likely due to doubt, as on harder questions children will have collected less net evidence and so will be expected to be more likely to reverse decisions in the absence of other influences. Clear evidence of the increasing adaptive value of social information use across childhood can be seen from the increase in the performance benefit due to social information (see **Figs. 4.3a and b**). The youngest children show no such increase in performance, whilst older children show a marked increase, particular on more difficult trials. For the oldest children, the probability that they get the second decision right having seen 8 out of the 10 informants supporting the correct decision is almost 40% greater than the probability they get the correct answer through individual investigation alone. This difference disappears on the easier trials; however, this is not because older children lose this advantage, it is simply because on these questions their individual performance is already so high that there is little room for further improvement due to social information (see **Fig. 4.3b**).

#### **4.5.4 - Asocial performance**

The results provide good evidence that the experiment worked as intended, with accuracy increasing both with dot ratio, but also with age (as shown by Halberda and Feigenson, 2008). This is intuitively plausible because dot ratio affects trial discriminability (which determines the relative difficulty of the trial) and because the ability to discriminate is likely to increase with age. Furthermore, all effects remaining in the model for the initial 'asocial' decision interacted with dot ratio, suggesting that our analysis was well able to account for variation in performance across trials and individuals. Of further interest are other effects for which I did not

find evidence. There was only very weak evidence for an interaction between area-correlation and ratio, which would have been expected had children used area covered by dots as a cue to the number of dots yielding better performance on area-correlated trials. I deliberately disrupted the relationship between number and total area to prevent this from being a reliable cue. However, it is still important to note that children were able to see past area covered and focus solely on number. Additionally, though perhaps less surprisingly, I also find convincing evidence that there is no gender difference in performance on this task. Finally, again unsurprisingly, children showed no bias for a particular side.

#### **4.5.5 - Conclusions**

In sum, the effect of asocial and social information on children's decision-making changes with age towards the adaptive decision-making mechanisms predicted by theory and observed in adults. Young children's judgements are indistinguishable from random behaviour unless they are presented with a total majority in which case they are likely to follow the social information, sticking or shifting with the consensus. Older children display a more nuanced pattern. They perform above chance, recall their previous decision and are biased in its favour even if the trial is extremely hard. They also make more use of the social information than younger children, switching and sticking strategically to increase their performance. Overall, the findings show that the mechanism for incorporating social information into decision-making is initially very blunt, only sensitive to overwhelming social signals. Across the course of early childhood, however, it increasingly responds to small

majorities, converging on those learning mechanisms observed in adults and predicted by theory.

## Chapter 5

# **An Empirical Case-Study of the Social Transmission of Oldowan Lithic Flaking Technology**

## **Abstract**

The simplest stone tools made by our hominin ancestors are sharp stone flakes that appear in the archaeological record from 2.5mya. However, the mechanisms by which the technology was transmitted remain unknown with suggestions that ape-like imitation could be sufficient, or alternatively that such tools imply more complex forms of social transmission such as teaching and language. Also unknown is why this technology persists with little alteration for around one million years. In this chapter I present an experiment investigating the efficacy with which different forms of social learning facilitate the transmission of the ability to manufacture flint flakes. The relevant skills were transmitted along chains of participants in five different conditions: (1) observation of flakes only, (2) observation of a tutor, (3) basic teaching from a tutor, (4) gestural teaching from a tutor, and (5) verbal teaching from a tutor. I find that more complex communication greatly assists the transmission of early stone knapping technology allowing participants to knap a greater proportion of their core and to make more viable flakes; however, observation of a tutor seems little better than simply observing pre-made flakes. I conclude that, whilst in the early Oldowan knapping may have been sustained by observational learning, the use of stone tools would have generated selection for more complex means of social transmission. I suggest that teaching, and perhaps a form of symbolic communication, evolved across the Oldowan. More complex technologies would have then appeared when hominin communication was sufficient to sustain them.



## 5.1 – Introduction

Animal bones marked by stone tools used for flesh removal appear in the archaeological record from 3.4mya (McPherron et al., 2010). From 2.52mya preserved stone tools show hominins were regularly making and using stone tools by striking cobble cores with a hammerstone to produce sharp flakes (Mode 1, Oldowan technocomplex; Roche et al., 1999; Semaw, 2000). This technology requires the selection of a core and hammerstone of appropriate size, shape and material, knowledge of the subtleties of flake manufacture (encompassing core and hammerstone positioning, target identification, striking angle and velocity, correction of knapping accidents, and maintenance of appropriate flaking angles throughout core reduction- cf. Delagnes and Roche, 2005) and efficient flake and core usage (e.g. butchering carcasses, scraping wood, cutting plants). This complexity supports the findings of toolmaking experimentation (Toth, 1987) in implying that Palaeolithic technology was acquired through learning and required considerable practice (Callahan, 1979; Roche, 2005). Early hominins removed more than 70 flakes from a single core, maintaining precise flaking angles and systematic patterns of flake detachment, including repairing damaged cores, which together implies some degree of planning and deployment of technological rules with sequential structure (Delagnes and Roche, 2005). The continuous existence and complexity of core-and-flake technologies, from 2.52mya, across Africa and Eurasia (Braun and Hovers, 2009), indicates that aspects of this technology were socially transmitted. Whilst the archaeological record can provide a detailed account of when and where technologies spread (Coward et al., 2008; Conolly et al., 2011), identifying the means of cultural transmission underlying a technology, when only the technology, and neither the

transmission nor transmitters, remains is a very difficult task (Shennan, 2011). Accordingly, the psychological mechanisms that underlie transmission of toolmaking technology remain unknown (Hovers, 2012). While the nature of Oldowan technological variation is subject to debate (Torre, 2011; Stout et al., 2010), that these industries persist for 770,000 years before the earliest Acheulean emerges, with no widely accepted explanation for this apparent stasis, highlights the need for an understanding of the mechanisms supporting stone knapping.

The relationship between Palaeolithic stone tool use and the evolution of human language and teaching (defined as active information donation, Fogarty et al., 2011) is unknown, yet highly debated (Gibson and Ingold, 1993; Ambrose, 2001; Byrne, 2005). One view is that the earliest form of stone tool-making in the Oldowan corresponds to a major development in hominin cognitive evolution (Hovers, 2012), for instance, indicative of instructive teaching or language (Bickerton, 2009). Alternatively, chimpanzee-like social learning through emulation (reproducing the object manipulations of others) and/or imitation (reproducing the motor patterns of others) has been argued as sufficient to transmit knapping technology (Wynn et al., 2011). Accordingly, estimates for the origin of language differ widely, from a few tens of thousands, to nearly two million, years (Belfer-Cohen and Goren-Inbar, 1994; Coolidge and Wynn, 2005).

Archaeological experiments with contemporary humans have provided valuable insights into cognitive and motor processes supporting lithic technology (Stout et al., 2000). Hence, toolmaking experiments can shed light on the above issues by establishing exactly what kinds of transmission mechanisms, verbal or otherwise, are

necessary for the long-term propagation of lithic technology. In this chapter I describe a large-scale, quantitative experimental investigation to test the power of alternative social learning mechanisms, notably observational learning (including imitation of actions and emulation of the results of actions), gestural teaching and verbal instruction, to transmit Oldowan stone knapping techniques along chains of learners, simulating successive generations. There is only one earlier example of such an approach (Ohnuma et al., 1997), which was limited to single transmission events and contrasted the roles of only speech and symbolic gestural communication in the acquisition of Levallois knapping techniques. In this chapter I rigorously measure the ability to make stone tools across successive transmission events, so as to address current issues concerning stasis and evolutionary change in the Oldowan technocomplex (Stout et al., 2010; Roche, 2005; de la Torre, 2011) as well as the likely roles of teaching, language and other factors in its transmission across generations.

## **5.2 – Methods**

### **5.2.1 - General Methods**

Across two weeks 184 participants learnt and taught others to make flint flakes using a granite hammerstone and flint core. I used a transmission chain design in which the first participant in a chain was taught by a skilled experimenter (either myself, or Natalie Uomini, both of whom have extensive practise in making flint flakes) and

subsequent participants were taught by the previous participant. Participants gained asocial information through access to the materials themselves. The social information was from a tutor and varied across five learning conditions detailed below. For each of the learning conditions four short chains ( $\leq 5$  participants long) and two long chains ( $\leq 10$  participants long) were carried out, totalling 30 chains across all conditions. Each participant was involved for ~90 minutes and was paid between £10 and £20 depending on their performance.

### **5.2.2 - Apparatus & Set-up**

The experiment used 2 tonnes of Brandon flint from a chalk quarry (Norfolk, UK), broken up into cores of roughly 1kg in weight. Helped by Katherine Meacham, Luke Rendell and Sally Street, I collected around 100 granite hammerstones, of a range of shapes and sizes from the coastline near Stonehaven, Scotland.

There were two knapping rooms, each of which contained a 4x4m square knapping area, the floor of which was covered in cardboard or black plastic sheeting, divided into two 2x4m sections by a 1m tall clear Perspex screen. In each section was a chair on which participants could sit and a large piece of Hessian that participants could use to protect their clothing whilst knapping. When only one participant was present they were free to use either section, but when a tutor and pupil were both present they each used one section. Participants were free to enter each other's sections during the pupil/tutor phases, but were only allowed to knap in their own section. Only one chain was carried out in each room at any given time, such that the only transmission of

information was from a single tutor to a single pupil and there was no horizontal transmission of information between multiple pupils. The screen ensured that flakes from each participant did not enter the other participant's section. Thus, it was clear who had produced any flakes found in each section. The screen also prevented flakes produced hitting another participant. Immediately to the side of the knapping area was a large pile of hammerstones from which participants were free to choose. For safety, all participants were required to wear a pair of safety glasses and latex coated cotton gloves. Additionally, breathing masks were provided for participants, in case they found the dust produced to be irritating. Two experimenters were present, at all times, sitting at a desk outside of the knapping area (which two experimenters were present varied: see the contributions section at the start of this thesis). A small number of flint cores were stored behind the desk and the experimenters chose cores from this supply at random for each participant.

### 5.2.3 - Procedure

Upon arrival, participants were briefed on the experimental procedure and given the opportunity to ask any questions. Participants then began the **introductory phase** of the experiment. Participants were provided with some pre-knapped flint flakes, some chamois leather and some sticks. They were given an information sheet containing limited information on the emergence of such technology in the archaeological record, descriptions of the tasks for which flakes were used, and informed that flakes were produced by striking pieces off a larger stone. They were then given 5 minutes to use the flakes to cut the leather and to sharpen the sticks. They were encouraged to try a

range of flakes to achieve an understanding of what properties made a useful (henceforth “viable”) flake and the provided flakes were deliberately of a variable quality. The introductory phase took part in a different room to the other phases of the experiment.

After this, the **pupil phase** began. Participants were given five minutes to practice making their own flint flakes. Additionally participants were provided with social information, the form of which varied depending on the learning condition, as detailed below, but included teaching by another participant if applicable (see **Section 5.2.4**).

Next, participants entered the **test phase**. They were instructed to make as many high quality flakes from the core as they could. They were not given an explicit time-limit, although the experimenter encouraged them to finish after 18 minutes and ended the phase if the participant took over 20 minutes.

If applicable (see **Section 5.2.4**), the participants next continued to the **tutor phase** where they provided social information to the next participant in the chain, just as they had experienced in their pupil phase. After this, participants were debriefed and were paid according to their performance (for details see **Section 5.2.5**).

In all phases of the experiment that involved knapping, participants were provided with a flint core and could choose a hammerstone. At the end of each phase that

involved production of flakes, participants were asked to separate out their flint into three categories; (i) what remained of the core, (ii) viable flakes and (iii) non-viable flakes. Flakes the participant selected as viable will henceforth be referred to as “selected”, whilst those they did not select as viable will be referred to as “non-selected”.

#### **5.2.4 – Learning Conditions**

The experiment involved 5 different learning conditions that dictated the form of the social information transmitted between individuals by placing limits on the ways in which the pupil and tutor could interact. The conditions were as follows:

**Reverse Engineering (“RE”)** - The pupil had access only to the flakes produced by their tutor, but no access to the tutor themselves. In this condition there was no teaching as the tutor was not present. Thus once participants had completed the test phase they proceeded immediately to debriefing. The flakes available to the pupil were those produced by the preceding participant in their test phase that the preceding participant had categorized as viable.

**Observational Learning (“OL”)** - The pupil was able to watch a tutor (the preceding pupil in the chain) making flakes, but no forms of direct interaction were permitted. As the tutor produced flakes they categorized them as viable or non-viable and the flakes were available for the pupil to examine should they wish. This condition allowed response facilitation as well as imitation and emulation (Byrne, 2009; Hoppitt

and Laland, 2008). Whilst response facilitation seems widespread, imitation and emulation seem much less so; particularly if imitation is limited to cases where the imitated behaviour is novel to the imitator, where it seems limited to the great apes (Byrne et al., 2011; Akins et al., 2002; Bates and Byrne, 2010; Byrne, 2009).

**Basic Teaching (“BT”)** – Teaching of the pupil by the tutor was permitted but was limited to a few simple forms. The permitted interactions were manual shaping (adjustment of the pupil’s hand position on their core and hammerstone via direct physical contact by the tutor), slowing of actions, and reorientation to allow the pupil a clear view. These forms of teaching were chosen as they are the forms of teaching for which there is anecdotal evidence in chimpanzees (Boesch, 1991, 1993) and hence can be regarded as an upper bound of what non-human primates are capable.

**Gestural Teaching (“GT”)** - Communication between the tutor and pupil was permitted but was limited to gestural (i.e., non-verbal) communication. This included, but was not limited to, mutual touching of tools, pointing, miming and nodding. This condition was chosen as there is evidence for great-ape gestural communication (Hobaiter and Byrne, 2011; Genty et al., 2009) and many have argued that verbal language originated from gestures (e.g., Sterelny, 2012). However, this does not mean I limited participant interactions to those gestures observed in non-human apes; all gestures were permitted, including, for example, nodding or head shaking to mean yes or no, neither of which are known outside humans.

**Verbal teaching (“VT”)** – All forms of communication between the tutor and pupil were permitted, including use of language.



In all teaching conditions the tutor was provided with their own flint core and hammerstone and could make their own flakes. Once flakes had been made the pupil was allowed to examine them.

### **5.2.5 – Participant Payment**

Participants were informed in advance of the payment scheme for the experiment, which varied by condition. In all conditions, participants were scored according to the number of viable flakes they were able to produce, divided by the initial mass of their core, during their test phase. This score was then translated into a payment (see below for what score was required for what payment). This score included any flakes that the experimenter considered viable, regardless of whether the participant had categorized them as such. This was to avoid the participant being motivated to categorise everything they produced as viable to maximise their payment. I chose this payment scheme as it reflects pressures on early hominin tool makers who lived in an environment where high-quality knapping material was sufficiently scarce that being able to produce many flakes from a given mass of knapping material would be a valuable skill (Diez-Martín et al., 2009; Braun et al., 2008). For the most basic condition (RE) this was the only influence on participant payment, such that:

$$\text{Score} = \frac{\text{number of viable flakes produced in test phase}}{\text{initial mass of core in test phase}}$$

In OL, participants were evaluated both their test phase and tutor phase performance; this was to motivate them to focus on their own performance during the tutor phase, instead of teaching the pupil. In this case, payment was calculated such that:

$$\text{Score} = 0.5 * \left( \frac{\text{number of viable flakes produced in test phase}}{\text{initial mass of core in test phase}} + \frac{\text{number of viable flakes produced in tutor phase}}{\text{initial mass of core in tutor phase}} \right)$$

In teaching conditions, participants were also evaluated on their student's subsequent test phase performance; this was to ensure tutors were motivated to teach effectively.

$$\text{Score} = 0.5 * \left( \frac{\text{number of viable flakes produced in test phase}}{\text{initial mass of core in test phase}} + \frac{\text{number of viable flakes produced by pupil in test phase}}{\text{initial mass of core in pupil's test phase}} \right)$$

The scores required for different payments are as follows:

Payment	£10	£12	£14	£16	£18	£20
Score required	0	5	10	15	20	25

### **5.2.6 - Recorded Variables**

Digital video cameras were used to record the entirety of the experiment (although video recording failed for one of the long chains in VT). Additionally, we recorded the initial weight of all the flint cores given to participants. Finally, at the end of each phase and for each participant we collected (i) what remained of the core, (ii) any selected material and (iii) any non-selected material, and placed each in a labelled bag.

## **5.3 – Coding**

### **5.3.1 – Individual flakes**

All pieces of flint greater than 2cm across were coded, totalling 6214 pieces. This lower limit of 2cm was judged the smallest a flake could be whilst retaining any value as a butchery tool (Braun et al., 2009; Key and Lycett, 2014). Analyses that found strong effects of condition were repeated with the lower limit on size set to 5cm as a test of the robustness of any findings. This did not change the findings and so I only present the results of the analyses containing data from all flakes over 2cm. Any flakes that had an edge deemed sharp enough to be of use were coded as viable, otherwise they were coded as non-viable. Prior to the full coding, a subset of 317 flakes were triple coded by myself (henceforth TM), Natalie Uomini (NU) and

Ignacio de la Torre (IT) to establish coder reliability. All three coders coded each piece of flint as viable or non-viable, and if viable, flakes were then rated on a 10-point scale of quality that took into account the efficiency with which the raw material had been used (non-viable flakes can be assumed to have a quality of 0). A latent variable analysis of flake viability was carried out to estimate the reliability of the viability coding decisions of each of the coders. The viability of each flake was modelled as a latent variable with a Bernoulli error structure (i.e. viable or non-viable). Additionally the viability ratings of each coder were modelled with a Bernoulli error structure and a logit link function. The linear predictors corresponding to coders' ratings took independent values for each value of the latent variable (i.e., viable or non-viable, allowing coder's to have different probabilities for false positives and false negatives). The only constraint placed upon the model was that all coders performed above chance, such that they had a >50% chance of identifying a flake correctly. This assumption implies that agreement between coders provides evidence concerning the actual viability of each flake and evidence that agreeing coders are more accurate than a dissenting coder. Thus the model used the coders' decisions to simultaneously estimate the viability of each flake and the accuracy of each coder. All three coders were estimated to have comparable levels of accuracy (estimated probabilities of accurate identification; TM = 0.81 [0.75, 0.87], NU = 0.89 [0.83, 0.94], IT = 0.82, [0.74, 0.88]). The imperfect viability coding by each of the three coders likely reflects the inherent difficulty in the coding decisions, as many flint fragments were of debatable value. The remaining flakes were coded by TM. In addition to viability, I also recorded flake cutting edge length, flake diameter and flake mass.

Based upon the 10-point quality ratings by the triple coders, a metric for flake quality was developed such that all flakes could be assigned a numerical quality rating that could be subject to analysis. Following Braun and Harris (2003), the metric began with:

$$quality = \text{flake cutting edge} / \text{flake mass}^{(1/3)}. \quad (5.1)$$

This scores flakes according to how much cutting edge they have, but the cube root function prevents larger flakes from being penalised by their large size (when scaled up in size, a three dimensional object's mass will increase by the scaling factor cubed). However, as a result, this formula does not take into account flake size. Size is clearly of relevance to flake quality, as excessively small flakes will be unusable and excessively large flakes will be wasteful of raw material. To include flake diameter the metric was extended to include a function of flake size such that:

$$quality = (\text{flake cutting edge} / \text{flake mass}^{(1/3)}) * f(\text{flake diameter}), \quad (5.2)$$

where  $f$  was an unknown function, with the constraint that  $f(x) \geq 0$ . To estimate the shape of  $f$  the quality ratings of the three triple coders were modelled with a binomial error structure (where  $n$  [i.e., the number of trials] was 10, as the ratings were on a 10 point scale). The probability of a success was logistically transformed into the positive continuous variable “quality”, that was modelled with formula 5.2. As the shape of

the diameter function was unknown, it was modelled as categorical with different values for each centimetre in diameter. Visual inspection of the values of this function strongly suggested a cumulative exponential function was appropriate (see **Figure 5.1**) and so the model was re-run with the function of flake diameter as a cumulative exponential distribution such that:

$$quality = (flake\ cutting\ edge / flake\ mass^{(1/3)}) * (1 - \exp(-\lambda * (flake\ diameter - offset))),$$

(5.3)

where  $\lambda$  is a positive continuous variable that sets the gradient of the cumulative exponential function and  $offset$  is the minimum possible diameter of a flake to have any quality whatsoever.  $offset$  was given a uniform prior ranging between 0 and 2 as flakes cannot be less than 0cm across and it was already decided that flakes over 2 could have some quality. The model estimates of these two parameters were:  $\lambda = 0.31$  [0.28, 0.35];  $offset = 1.81$ , [1.69, 1.90]. The posterior distribution for  $offset$  sat comfortably within the interval specified by the prior, suggesting that it was an appropriate prior distribution. Given this, the final flake quality metric is:

$$quality = (flake\ cutting\ edge / flake\ mass^{(1/3)}) * (1 - \exp(-0.31 * (flake\ diameter - 1.81))).$$

(5.4)

This function rewarded flakes for a high cutting edge length and penalised flakes for being excessively small (see **Figure 5.1**). Around a size of 2cm flakes were very heavily penalised; however, the effect of flake diameter flattens above 6cm such that further increases in size do not greatly increase quality. It is of note that the diameter function does not penalise flakes for being excessively large. This is presumably because most flakes produced by participants were small and so very few flakes were large enough to merit penalisation for excessive size.

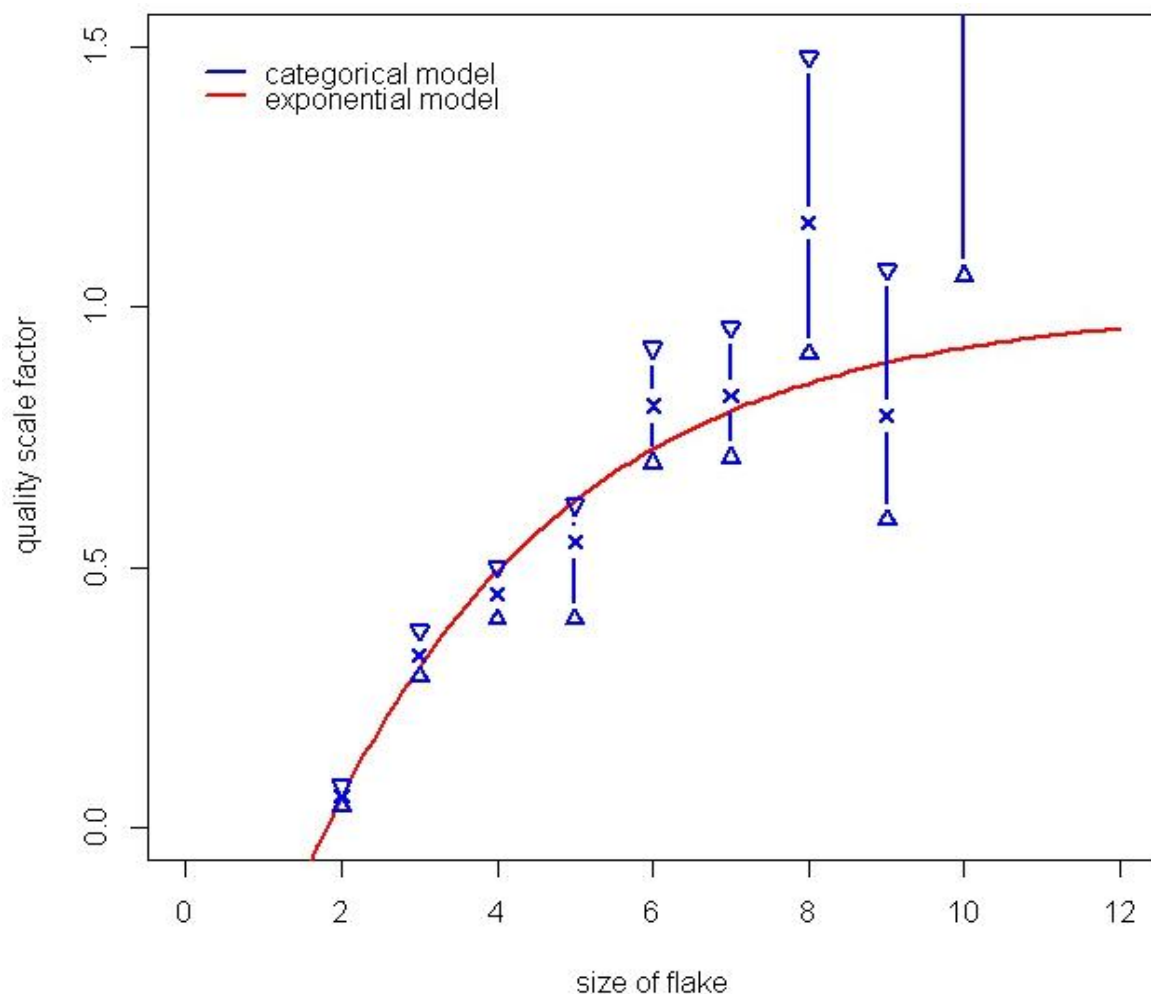


Figure 5.1

Estimation of the  $f(\text{flake diameter})$  function. The blue intervals show the medians and 95% central credible interval value of this function when it was modelled as categorical with different values for every cm of flake diameter. The intervals widen with increasing flake size as most flakes were small and so there is less data for larger flakes. The red line shows the

shape of formula 5.4 which modelled following visual inspection of the categorical model results and was subsequently used to evaluate the quality of individual flakes. That the red line fits the blue intervals well suggests that formula 5.4 is suitable for describing flake quality.

### 5.3.2 – Videos

The participants' behaviour, as video recorded at all points in the pupil, test and tutor phases, was coded as falling into one of the following categories:

**Knapping** - where the participant directs their attention toward their own core and hammerstone with the apparent intention of making flakes for their own ends. This is not limited to the act of striking the core and would also include examination of the core and hammerstone in preparation for striking.

**Observing** - where the participant directs their attention to the other participant or the flakes of the other participant, using them as a source of information. This is not limited to passive observations, and, in the conditions that permit this, would include active attempts to extract information from the other participant.

**Teaching** - where the participant seemingly aims to pass on information to the other participant. This would include instances where one participant strikes their core as a demonstration for the other participant.

**Choosing** - where the participant directs their attention to flakes that they have produced and considers the quality or nature of them. If the participant proceeds to try to knap the flake this no longer counts as choosing and instead counts as knapping.



Other - any behaviour that does not fit into the above categories.

Coders marked the points in time at which participants' behaviour changed from one of the above categories to another and so all points between these intervals can be categorised. Additionally, the time of every strike of the core with the hammerstone was recorded. As a test of coding accuracy, ten participants were randomly chosen (2 from each condition, 10% of all participants) and their videos were coded by both TM and Ronan Kearney (RK). I modelled the absolute magnitude of the disagreement between the two coders concerning total time spent knapping and the total number of times each participant hit their core with their hammerstone as these variables were used in further analyses. In the case of time spent knapping I used a gamma error structure and the expected difference is 20.4s, [14.0, 31.2]. As a proportion of the average time for which participants were present this is 0.04, [0.03, 0.07] which is a very low proportion of disagreement. In the case of total hits, I used a Poisson error structure and the expected disagreement is 7.7 hits [6.7, 8.8], as a proportion of the average number of times each participant hit the core with their hammerstone this is 0.04, [0.04, 0.05]. Given this high level of agreement, RK went on to code all the remaining videos.

### **5.3.3 – Language**

Whilst coding the videos as described above, RK also transcribed everything that was said by participants in the verbal teaching condition. This was then coded by TM as

follows. Initially, each transcript was split into utterances - where an utterance was defined as a single, continuous stretch of verbal communication by a single participant. Thus an utterance ends with a pause or when the other participant says something. Each utterance was scored according to the following categories which are not mutually exclusive, in that a single utterance could (in theory) score positively for every category:

1. Said by the tutor – utterances said by the tutor.
2. Teaching – utterances transmitting knapping-relevant information to the other participant (this could be from the pupil to the tutor), e.g., “You want to rest the flint core on your left leg” which transfers knowledge of how to hold the core.
3. Feedback - utterances giving feedback on performance, in terms of encouraging good behaviour or vice-versa. Feedback is a type of teaching, so any utterance that is categorized as feedback will also be categorized as teaching. e.g., “So that's the sort of thing you want to, that's brilliant”.
4. Confirmation of understanding - utterances with the apparent purpose of confirming that the speaker had understood something. Most instances of the word “yes” were coded in this category. e.g., “Ok, of course”, but not “So you're always trying to hit above a ridge then?” which would be coded as a request for information (category 7).
5. Watch this - utterances directing attention to the speaker in order to demonstrate something. e.g. “just...” followed by the speaker knapping.

6. This/that - utterances using words such as "this" or "that" to indicate objects or locations. e.g. "That one's no good, is it?"
7. Requesting Information - utterances requesting knapping relevant information. e.g. "So you're always trying to hit above a ridge then?", which requests information on where to hit.
8. Conveying uncertainty - utterances including an expression of uncertainty. e.g. "Maybe that bit's kind of hanging over and there's kind of an under-hang, try that". The uncertainty in this expression is conveyed with the terms "maybe", "kind of" and "try that".
9. Abstract - utterances using abstract descriptions that give general information not specific to a single case. e.g. "Find an edge, do you have an edge with black stuff on the other side as well?" This describes the general procedure for identifying an edge without cortex, and can be contrasted with "Emm, this is probably going to be your hit" where a participant simply points out a specific point with no generalizable information.
10. Correct – information in the utterance is factually correct.
11. Incorrect – information in the utterance is factually incorrect.

In addition to the above categories, the topic of the utterances (as opposed to their nature/purpose) was also categorized according to the following topics:

1. knapping (a broad category)
2. knapping site

3. platform edge
4. platform angle
5. ridge
6. how to hit
7. how to hold
8. hammerstones
9. cortex
10. choosing flakes (a broad category)
11. size of flakes
12. cutting edge of flakes
13. safety whilst knapping

As with the previous categories, the topics are not mutually exclusive. Additionally, topics 1 and 10 (knapping and choosing flakes) are very broad, with the other topics falling as sub-topics within these. For example, the topic “platform edge” is a sub-topic within “knapping” as by talking about the platform edge you are also talking about knapping.

## 5.4 – Analyses

### 5.4.1 – General Analyses

All analyses were carried out in OpenBUGS v3.2.1 using MCMC methods and all quoted values are the median and 95% central credible interval from a sample of at least 3000 iterations, following a sufficient burn-in period and thinning to remove autocorrelations. Unless otherwise stated analyses are based on data from the participants' test phases.

### 5.4.2 – Individual flakes

I analysed **flake mass, flake diameter, flake cutting-edge length and flake quality**. In all cases the variable was modelled with a log-normal error structure (censored below 2cm in the case of diameter) with independent effects of condition, a linear effect of position along the chain that interacted with independent effects of condition, a linear effect of core mass and random individual level effects. In the cases of flake mass and diameter these analyses were repeated using a data set containing all the flakes (n=6214) and the following subsets of flakes; (i) flakes coded as viable (n=3946), (ii) flakes coded as non-viable (n=2268), (iii) selected flakes (n=3007), and (iv) non-selected flakes (n=3207). In the case of flake cutting-edge length and flake

quality, the analysis was carried out only on the subset of viable flakes as, by definition, only these flakes had non-zero values for these variables.

### 5.4.3 – Aggregate flakes

I analysed **the total number of flakes and the number of viable flakes, non-viable flakes, selected flakes and non-selected flakes** that each participant produced with a Poisson error structure. I also analysed **the proportion of flakes that were viable and the proportion of flakes that were selected** using a binomial error structure. In this case the total number of flakes produced was used as the number of trials and the number of viable/selected flakes was the number of successes. The proportions of flakes that were non-viable or not selected were not analysed as they are the inverse of the proportion of flakes that are viable and selected respectively. Using a gamma error structure I also analysed **the sum of the cutting-edge length, the sum of the mass and the sum of the quality of all flakes** produced by participants. All of these models used a logarithmic link function, except for the binomial models that used a logistic link function, and the linear predictors contained categorical effects of condition that interacted with a linear effect of position along the chain and a linear effect of core mass. Individual-level effects were not included as each individual only contributed a single data point to each analysis.

#### 5.4.4 – Flake identification

I analysed the **probability that a viable flake was selected, a non-viable flake was not selected, a selected flake was viable and that a non-selected flake was non-viable** with a Bernoulli error structure and a logistic link function. In all four cases the linear predictor contained categorical effects of condition that interacted with a linear effect of position and random individual level effects. To investigate what led participants to select particular flakes, I also analysed **whether a flake was selected** by participants with a Bernoulli error structure and logistic link function. The linear predictor contained a categorical effect of condition that interacted with linear effects of flake quality metric, flake mass, flake cutting edge and flake diameter and random individual-level effects.

#### 5.4.5 – Core remaining

Using a hurdle model, I analysed **the proportion (by mass) of the participants' cores remaining** after knapping. First the analysis modelled whether a participant had any of their core remaining at all with a Bernoulli error structure and logistic link function, then in the cases where there was some core left, it modelled the proportion left with a beta error structure and logistic link function. These two elements could then be combined to produce an estimate of the expected core remaining, such that

$$p(left) = p(any\ left) * p(left/some\ left), \quad (5.5)$$

where  $p(left)$  is the expected proportion of core remaining,  $p(any\ left)$  is the probability that any core remains and  $p(left/some\ left)$  is the expected proportion of core remaining given that some remains. In both parts of the model the linear predictor contained categorical effects of condition that interacted with a linear effect of position along the chain. Individual level effects were not included as each individual contributed only a single data point to each analysis.

#### 5.4.6 - Videos

I modelled **the number of hits per minute spent knapping, the number of flakes produced per minute** (both for the case of all flakes and for just viable flakes) with a lognormal error structure (no link function required), and **the probability each hit produces a flake** (both for all flakes and viable flakes) with a binomial error structure and logistic link function. The linear predictor contained categorical effects of condition that interacted with a linear effect of position. There were no effects of core mass as it was deemed implausible that this would have an effect on the variables investigated. I also investigated how the behaviour of the tutor and pupil participants changed across conditions, along chains and contingent on tutor quality. I modelled **how much time in the their tutor phase tutors spent knapping, observing, teaching, choosing and how many times tutors hit their core with their hammerstone, how much time in their pupil phase pupils spent knapping,**



**observing and choosing and how many times pupils hit their core with their hammerstone** using a Poisson error structure and logarithmic link function. The linear predictors contained categorical effects of condition that interacted with linear effects of position along chains and the number of viable flakes the tutor had produced in their previous test phase, which is used as a measure of tutor quality. Effects of core mass were not included as it was deemed implausible that this would have an effect. Data on how tutors spent their time was taken from chains that allowed teaching (BT, GT and VT), as only these conditions involved teaching. Data on how pupils spent their time was taken from chains that had a tutor present (OL, BT, GT and VT).

#### 5.4.7 – Language

**The total number of utterances** said was analysed with a Poisson error structure. The model incorporated chain length with a function that set a baseline number of utterances, an initial deviation from this number and a rate parameter that set the rate at which the value approached the baseline from the initial value. The shape of the function was that of a cumulative exponential function. The model included a random effect of repeat for the initial value and did not need to include condition as data was taken from a single condition, as only the verbal teaching condition allowed language. I also analysed **the probability a given utterance satisfied each of the above categories or covered each of the above topics** with Bernoulli error structures and logistic link functions. The linear predictor used the same function as the model for the total number of utterances. I also investigated whether different topics were

transmitted with greater accuracy by modelling **whether an utterance was scored as correct or incorrect** with a Bernoulli error structure and logistic link function. The linear predictor contained categorical effects of all the topics (other than knapping and choosing flakes as the sub-topics were included instead). Finally, I also investigated whether use of language could tell us anything about subsequent performance in the test phase. To this end I modelled **the number of viable flakes a participant produced** with a Poisson error structure and logarithmic link function. The linear predictor contained effects of how many teaching utterances the participant's tutor had made and the proportion of teaching utterances scored as right or wrong that were scored as right.

## 5.5 – Results

### 5.5.1 – General Results

Most of the results I shall present here fall into two categories - those concerning condition and others concerning position along the chain. Condition effects compare the effects of different forms of social transmission (i.e. verbal teaching versus gestural teaching) whilst position effects concern changes in performance along chains, but within a particular condition. Throughout this section, strong evidence for an effect corresponds to a 95% central credible interval which excludes 0, weak evidence corresponds to cases where the 95% central credible interval includes 0, but the 90% central credible interval does not, no evidence corresponds to an interval

centred around 0. Unless otherwise stated, the estimated values and their contrasts that are quoted are on the scale of the units of the value being analysed (e.g., units of mass for analyses of the mass of flakes), whereas the gradients of linear effects are quoted on the scale of the linear predictor. Unless otherwise stated, for the model estimates for each variable by condition see **Table 5.1**, for model estimates for contrasts between conditions see **Table 5.2**, for model estimates for the rate of change along chains see **Table 5.3**, and for model estimates for the effects of core mass see **Table 5.4**. Unless otherwise stated, all figures shows median model estimates with associated 95% central credible intervals.

### 5.5.2 – Individual Flakes

In general, there were few differences between conditions, or changes along chains, in the properties of individual flakes. Where evidence for differences between conditions is found, it is not consistent across multiple conditions in that if there is strong evidence that a measure is greater with VT than with BT, there is typically not evidence that it is greater with VT than with OL or RE. Specifically:

There was strong evidence that the **mass of flakes** produced with BT was greater than with RE or OL. There is no evidence that flake mass changed along chains, or that core mass had an effect on flake mass.

There is strong evidence that **viable flake mass** was greater with BT than with OL (see **Fig. 5.2**). There is no evidence that flake mass changed along chains, or that core mass had an effect on viable flake mass.

There is no evidence **non-viable flake mass** changed across conditions. There is no evidence that non-viable flake mass changed along chains, or that core mass had an effect.

There is strong evidence that **selected flake mass** was greater with GT than with VT or RE. There is no evidence of an effect of position along the chain in any condition, although there was a positive effect of core mass on selected flake mass.

There is no evidence that **non-selected flake mass** changed across conditions. There were no effects of position along the chain and no effect of core mass.

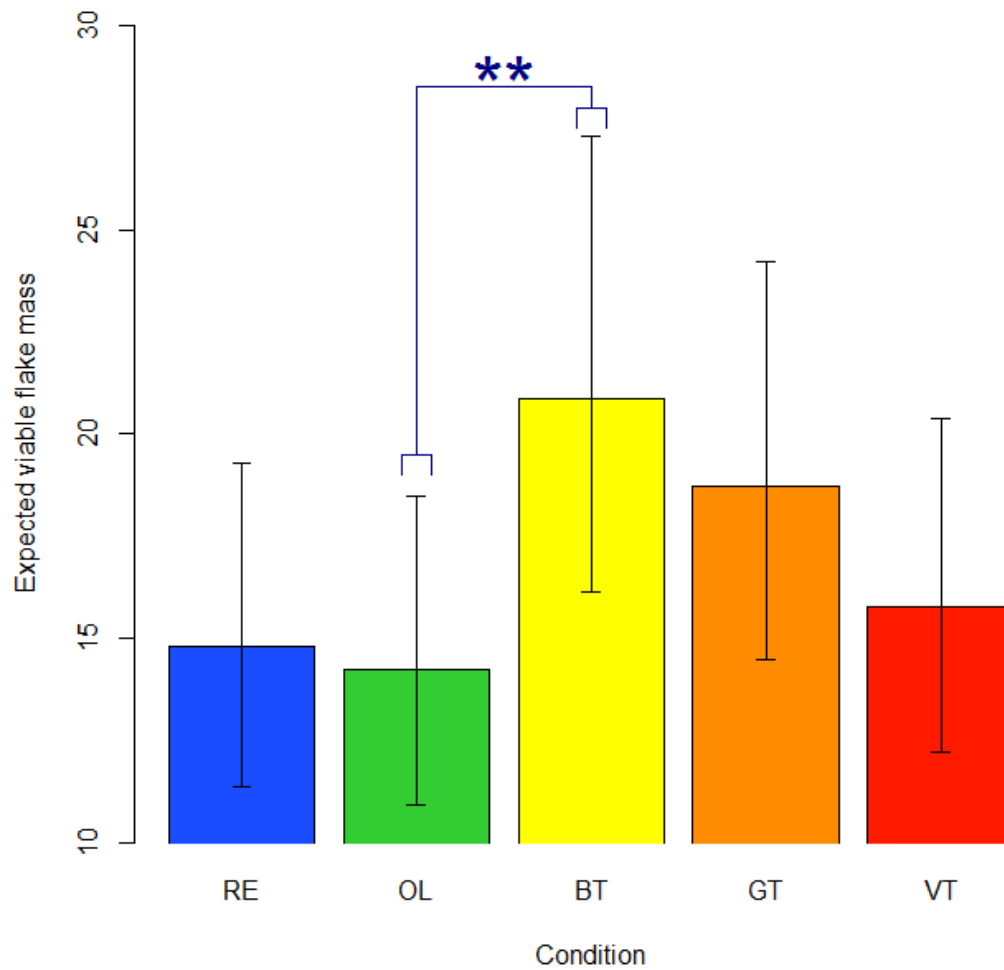


Figure 5.2

The expected mass of viable flakes shows no clear trend across conditions, although it is higher with Basic Teaching than with Observational Learning. The brackets indicate contrasts for which there is evidence of a difference, single asterisks correspond to weak evidence, double asterisks to strong evidence. The height of each bar is the median model estimate, whilst the error bars show the 95% central credible interval.

There is strong evidence that **the diameter of all flakes** was greater with BT than with RE or OL and greater with GT than with RE. There is no evidence that position in the chain affected flake diameter in any condition, or that core mass had an effect.

There is strong evidence that **the diameter of viable flakes** was greater with BT than with RE or OL. There is no evidence of an effect of chain position in any of the conditions and no evidence of an effect of core mass.

There is strong evidence that **the diameter of non-viable flakes** was greater with BT than with RE or OL. There is no evidence that position along the chain had an effect on non-viable flake diameter or that core mass had an effect.

There is strong evidence that **the diameter of selected flakes** was greater with GT than with RE or VT and greater with BT than with VT. There is no evidence of an effect of chain position in any of the conditions, but there is evidence of a positive effect of core mass on the diameter of selected flakes.

There is little evidence for differences in **the diameter of non-selected flakes** across conditions. Neither is there evidence for an effect of chain position in any conditions or an effect of core mass.

There is no evidence of differences in **flake cutting edge** across conditions and there is no evidence of a difference between conditions. There is no evidence of an effect of chain position or of core mass.

There is no evidence of differences in **flake quality** across conditions (see **Fig. 5.3**).

There is no evidence for an effect of chain position, or for an effect of core mass.

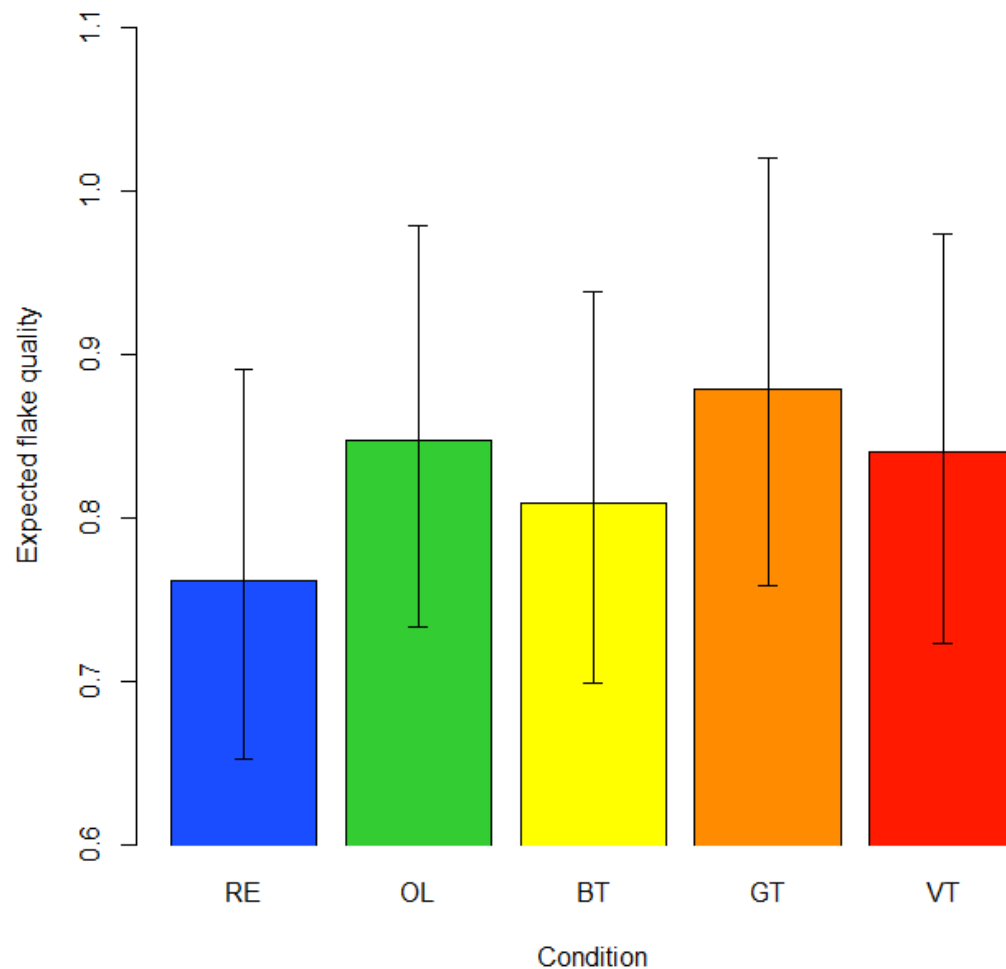


Figure 5.3

The expected quality of an individual flake shows no evidence of changing across conditions. The height of each bar is the median model estimate, whilst the error bars show the 95% central credible interval.

### 5.5.3 – Aggregate flakes

In contrast to the analyses of individual flakes, there is often strong evidence for a difference between conditions in aggregate measures of flakes. Furthermore these differences are consistent across multiple conditions in that if there is strong evidence

that a measure is greater with VT than with BT, it is typically also greater with VT than with OL or RE. Specifically:

There is strong evidence that **the total number of flakes** was greater with VT than with any other condition and greater with OL than with BT. The number of flakes produced increased along chains with RE, BT, and potentially GT, but there was no strong evidence for a change along chains with OL and a negative effect with VT. There was a positive effect of core mass on the total number of flakes produced.

There is strong evidence that **the number of viable flakes** was greater with VT than with RE and weak evidence that it was greater with VT than OL and greater with GT than with RE (see **Fig. 5.4**). Performance decreased along chains with VT, but increased with RE (see **Fig. 5.5**). There is no evidence of a change in performance along chains in the other conditions, but there is strong evidence for a positive effect of core mass. There is strong evidence that the **proportion of flakes that are viable** was greater with BT, GT, or VT than with RE or OL (see **Fig. 5.6**). There is also evidence that the increase in the proportion of flakes that are viable between OL and BT was greater than the increase between any other adjacent conditions. The proportion of flakes that are viable decreased along chains in BT, GT and VT (see **Fig. 5.7**).

There is strong evidence that the **number of non-viable flakes** was less with GT and BT than with RE and weak evidence that it was less with BT and GT than with OL. The number of non-viable flakes produced increased along chains with BT and GT, but decreased with OL, and there was a positive effect of core mass.

There is strong evidence that the **number of selected flakes** was greater with VT than with GT, OL or RE and weak evidence that it was greater with VT than with BT. The



number of selected flakes increased along chains in RE and decreased with VT, but there was no change in any other condition. There was a negative effect of core mass. There is strong evidence that the **proportion of flakes that were selected** was greater with VT and BT than with RE, OL or GT. There was a decrease in the proportion of flakes selected along chains with BT and VT.

There is strong evidence that the **number of non-selected flakes** was lower with BT than with OL. The number of non-selected flakes increased along chains with RE, BT and GT, but decreased with OL. There was a positive effect of core mass.

There is strong evidence that the **total cutting edge** was greater with VT than with RE, OL or BT and weak evidence it was greater with GT than with RE. There was no effect of chain position and no effect of core mass.

There is weak evidence that the **total mass of flakes produced** was less with RE than with VT GT or BT. There is no evidence for an effect of position along chains or of core mass.

There is strong evidence that the **total quality of flakes produced** was greater with VT than with RE, and weak evidence that it is greater with VT than with OL or BT and greater with GT than RE (see **Fig. 5.8**). There is no evidence of an effect of position along the chain in any condition, and no evidence of an effect of core mass.

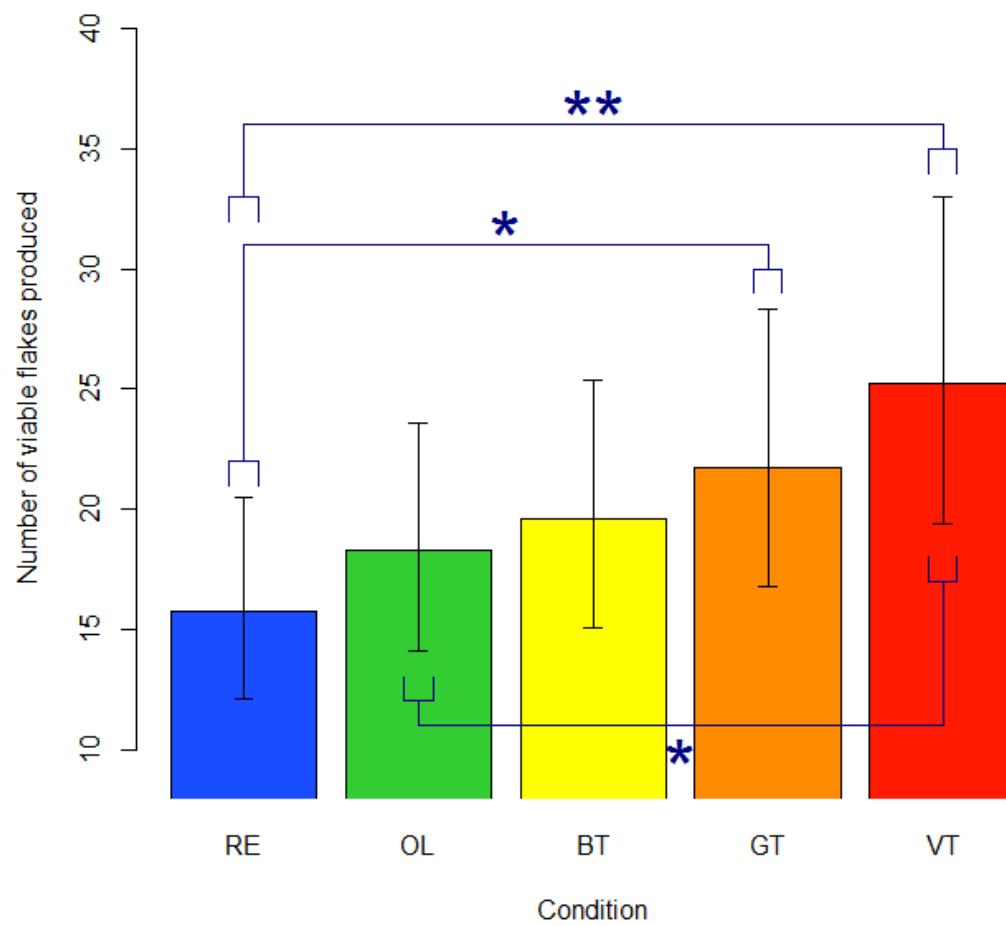


Figure 5.4

The number of viable flakes produced across conditions. The brackets indicate contrasts for which there is evidence of a difference, single asterisks correspond to weak evidence, double asterisks to strong evidence. The height of each bar is the median model estimate, whilst the error bars show the 95% central credible interval.

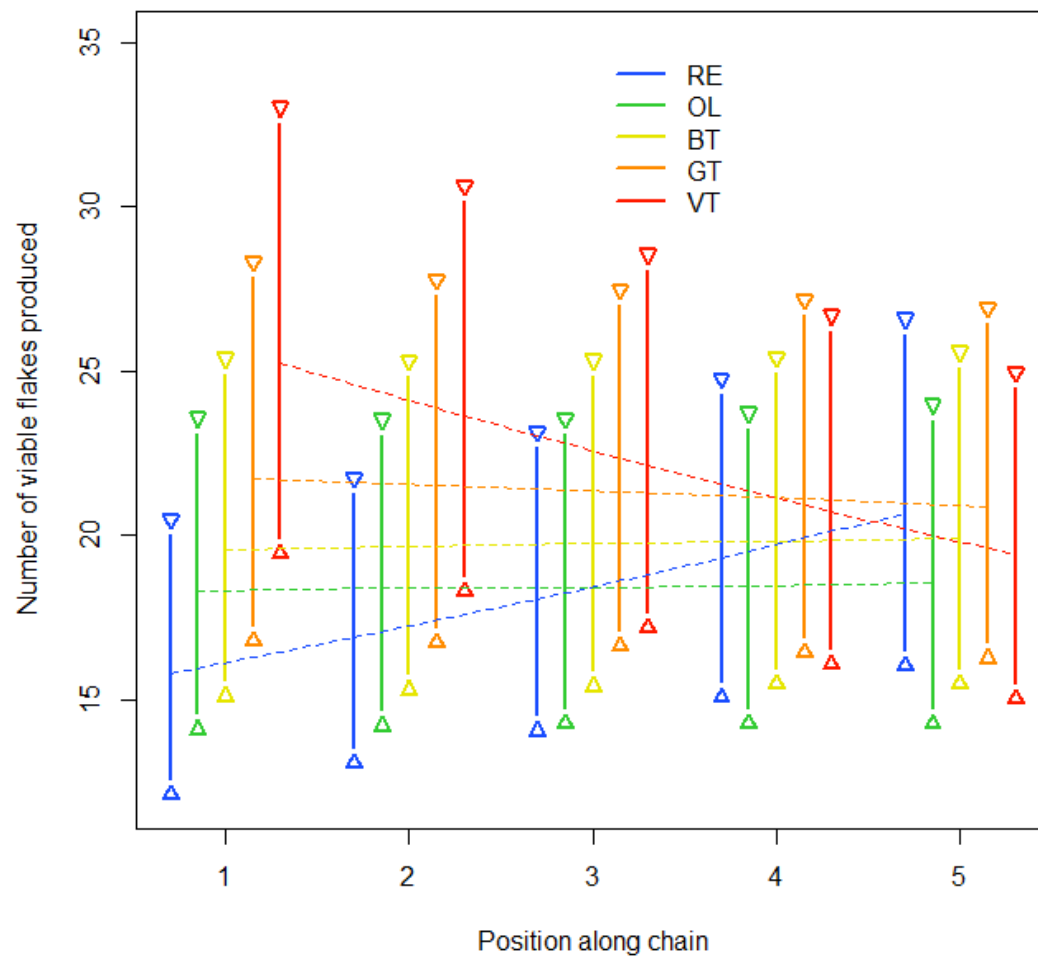


Figure 5.5

Differences between conditions in the number of viable flakes produced by participants disappear along chains. The solid lines represent the 95% central credible intervals for estimated values at the first five points in chains for each condition, whilst the dashed lines link their medians to show the trend along chains. In this case there is an increase in performance in the case of chains with only Reverse Engineering.

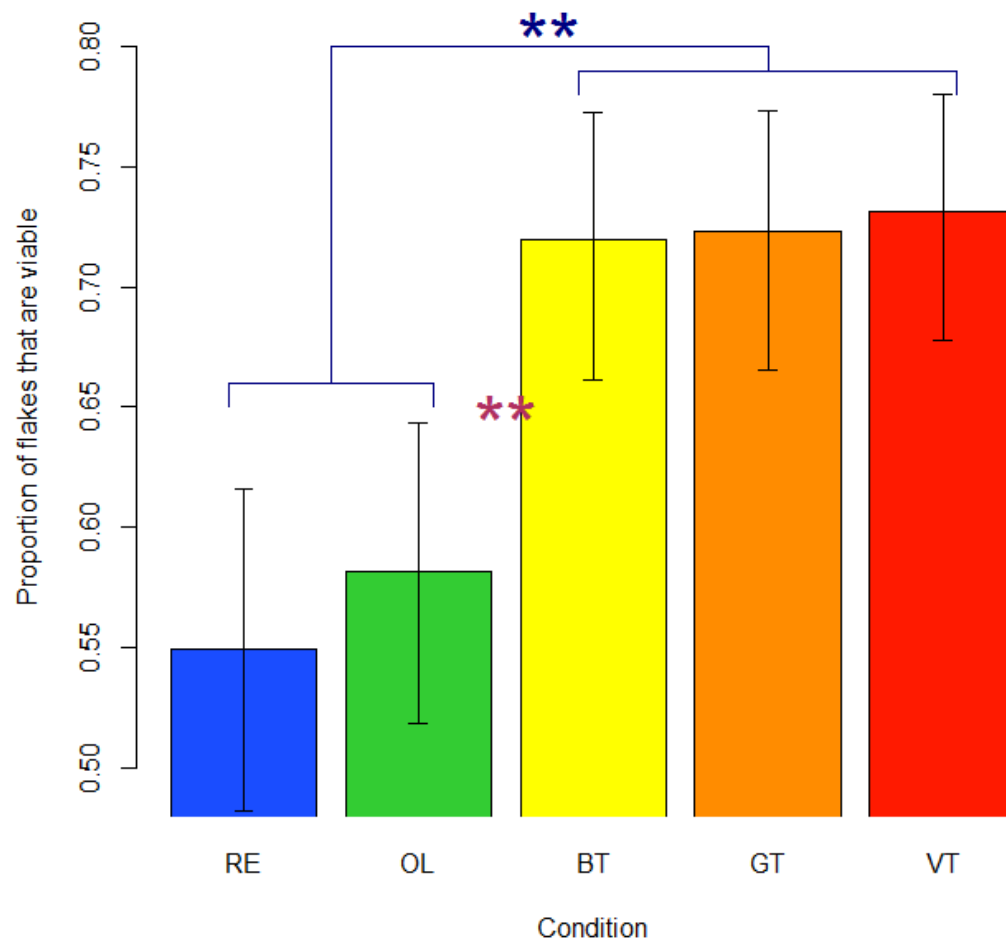


Figure 5.6

The proportion of flakes produced by participants that are viable is greater with teaching (Basic, Gestural or Verbal Teaching) than without (Reverse Engineering and Observational Learning). The brackets indicate contrasts for which there is evidence of a difference, single asterisks correspond to weak evidence, double asterisks to strong evidence. The double red asterisk indicates that there is strong evidence that the increase in performance between Observational Learning and Basic Teaching is greater than the increase in performance between all other adjacent pairs of conditions. The height of each bar is the median model estimate, whilst the error bars show the 95% central credible interval.

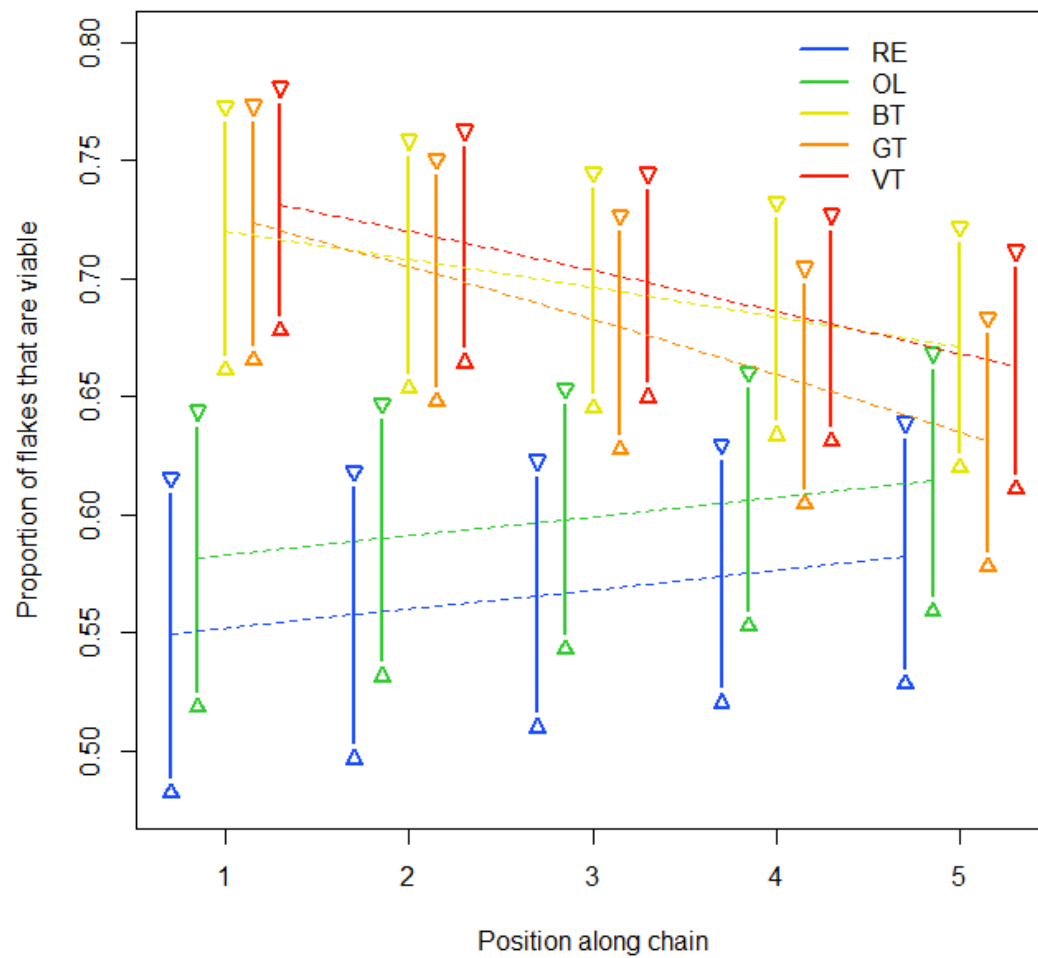


Figure 5.7

The differences between conditions in the proportion of flakes that are viable disappear along chains. There is no evidence for a change along chains in conditions without teaching. The solid lines represent the 95% central credible intervals for estimated values at the first five points in chains for each condition, whilst the dashed lines link their medians to show the trend along chains.

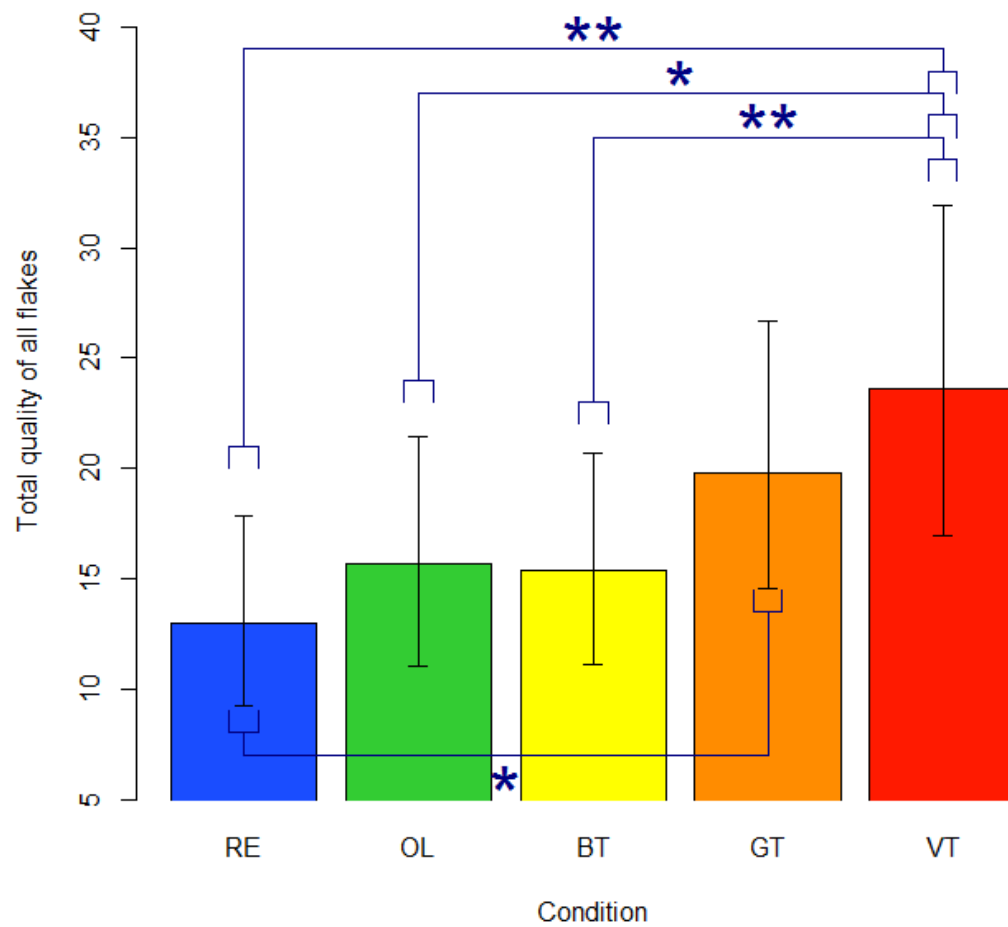


Figure 5.8

The total quality of all flakes produced by participants is greater with Verbal Teaching than with Reverse Engineering, Observational Learning and Basic Teaching and greater with Gestural Teaching than with Reverse Engineering. The brackets indicate contrasts for which there is evidence of a difference, single asterisks correspond to weak evidence, double asterisks to strong evidence. The height of each bar is the median model estimate, whilst the error bars show the 95% central credible interval.

#### 5.5.4 – Flake Identification

There is no evidence that **the probability a viable flake was selected** differed between conditions or changed along chains. There is no evidence that **the probability a non-viable flake was not selected** differed between conditions or changed along chains. There is strong evidence that **the probability that a selected flake is viable** was greater with VT or GT than with RE (see **Fig. 5.9**). There is no evidence of an effect of chain position in any condition on the probability a selected flake is viable. There is strong evidence that **the probability a non-selected flake is non-viable** is greater with OL than with GT or VT (see **Fig. 5.10**). There is no evidence of an effect of chain position on the probability a non-selected flake is non-viable.

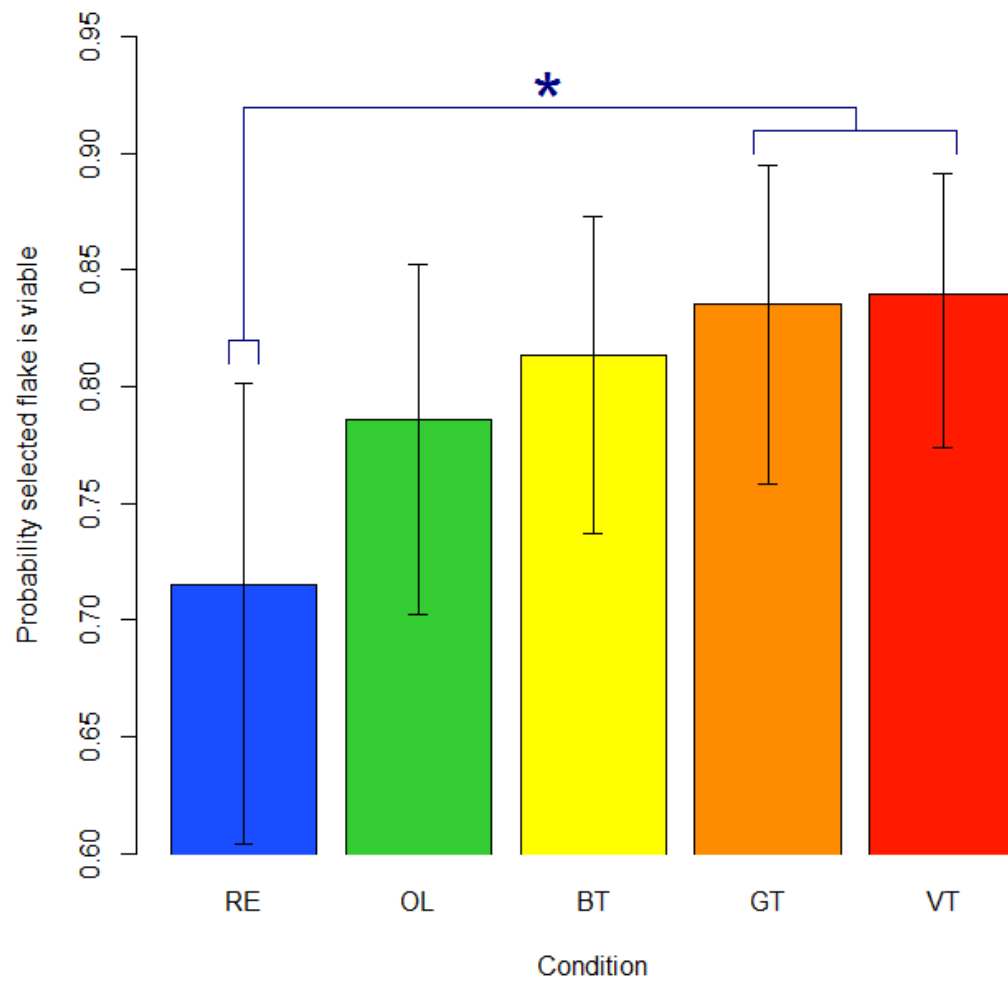


Figure 5.9

The probability that a flake selected by the participant is actually viable is greater with Verbal or Gestural Teaching than with Reverse Engineering. The brackets indicate contrasts for which there is evidence of a difference, single asterisks correspond to weak evidence, double asterisks to strong evidence. The height of each bar is the median model estimate, whilst the error bars show the 95% central credible interval.



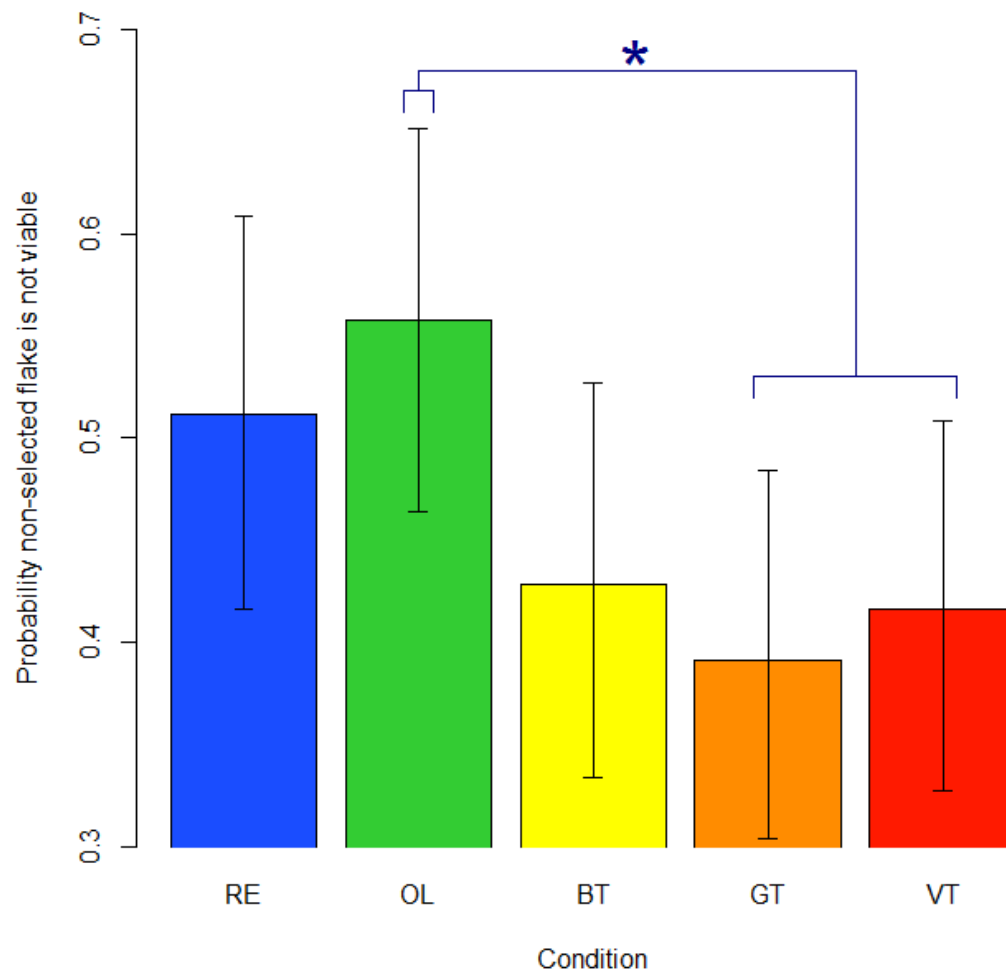


Figure 5.10

The probability that a flake discarded by the participant is actually non-viable is less with Verbal or Gestural Teaching than with Observational Learning. The brackets indicate contrasts for which there is evidence of a difference, single asterisks correspond to weak evidence, double asterisks to strong evidence. The height of each bar is the median model estimate, whilst the error bars show the 95% central credible interval.

### 5.5.5 – Core Remaining

There is strong evidence that the **proportion of the core left** was less with VT than with RE, and weak evidence that it was less with VT than with OL (see **Fig. 5.11**). Due to the nature of the model it is difficult to present an estimate of the change along chain lengths, to estimate this I present the effect of position along the chain on the proportion remaining given that some core remains (this seems reasonable as in only 9 cases was there no core left). There is no strong evidence of an effect of position along the chain on the proportion of core remaining in any condition. There is strong evidence of a negative effect of core mass on the proportion of core remaining.

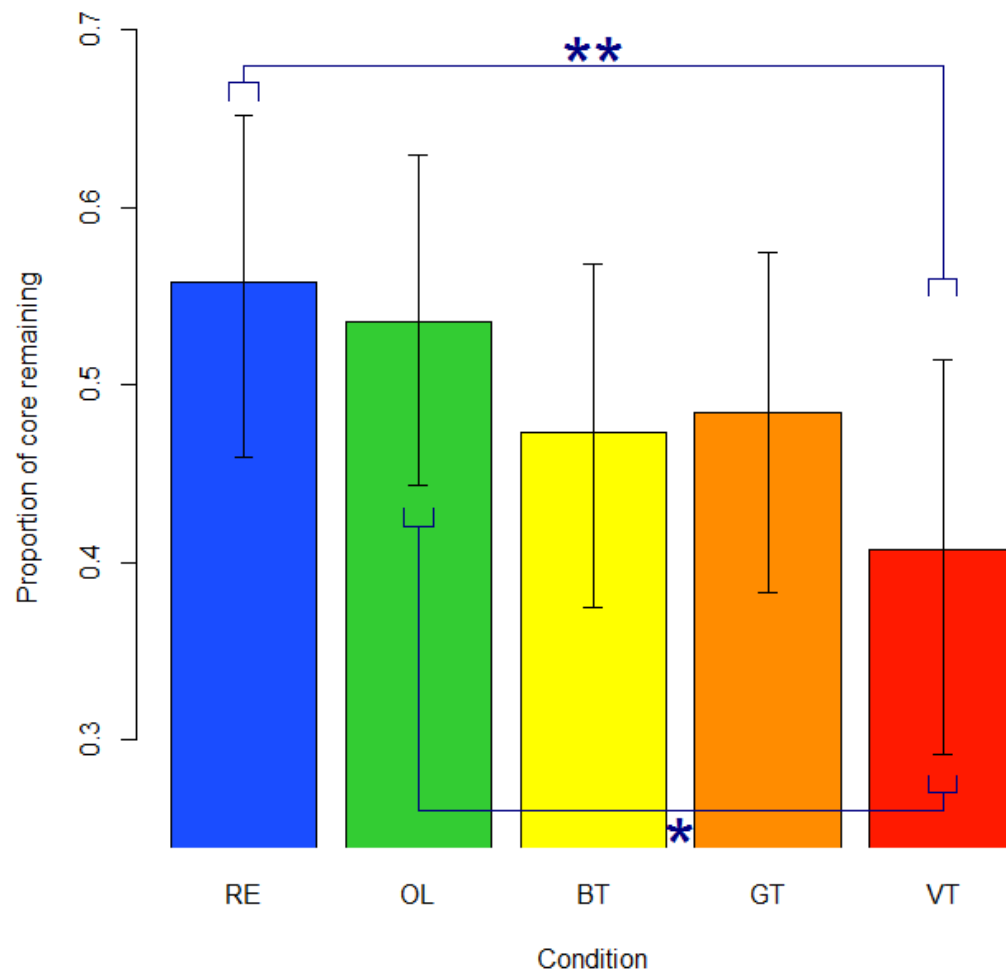


Figure 5.11

The proportion of the participant's core remaining unknapped at the end of the test phase is lower with Verbal Teaching than with Observational Learning or Reverse Engineering. The brackets indicate contrasts for which there is evidence of a difference, single asterisks correspond to weak evidence, double asterisks to strong evidence. The height of each bar is the median model estimate, whilst the error bars show the 95% central credible interval.

### 5.5.6 - Videos

There is strong evidence that the **number of hits per minute spent knapping** was greater with RE than VT (see **Fig. 5.12**). There is strong evidence that this increased along chains with VT, although not in any of the other conditions (see **Fig. 5.13**). There is strong evidence that the **number of viable flakes per minute spent knapping** was greater with VT than with RE or OL (see **Fig. 5.14**). There is no evidence that chain position had an effect on the number of viable flakes produced per minute spent knapping. There is strong evidence that the **probability each hit produces a viable flake** is greater with VT than with BT, OL or RE, greater with GT or BT than with RE. There is also weak evidence that it is greater with GT than with OL (see **Fig. 5.15**). Chain position had a negative effect on the probability each hit produces a viable flake with VT, GT and OL, there is also weak evidence of a negative effect with BT, but no evidence of such an effect with RE (see **Fig. 5.16**).

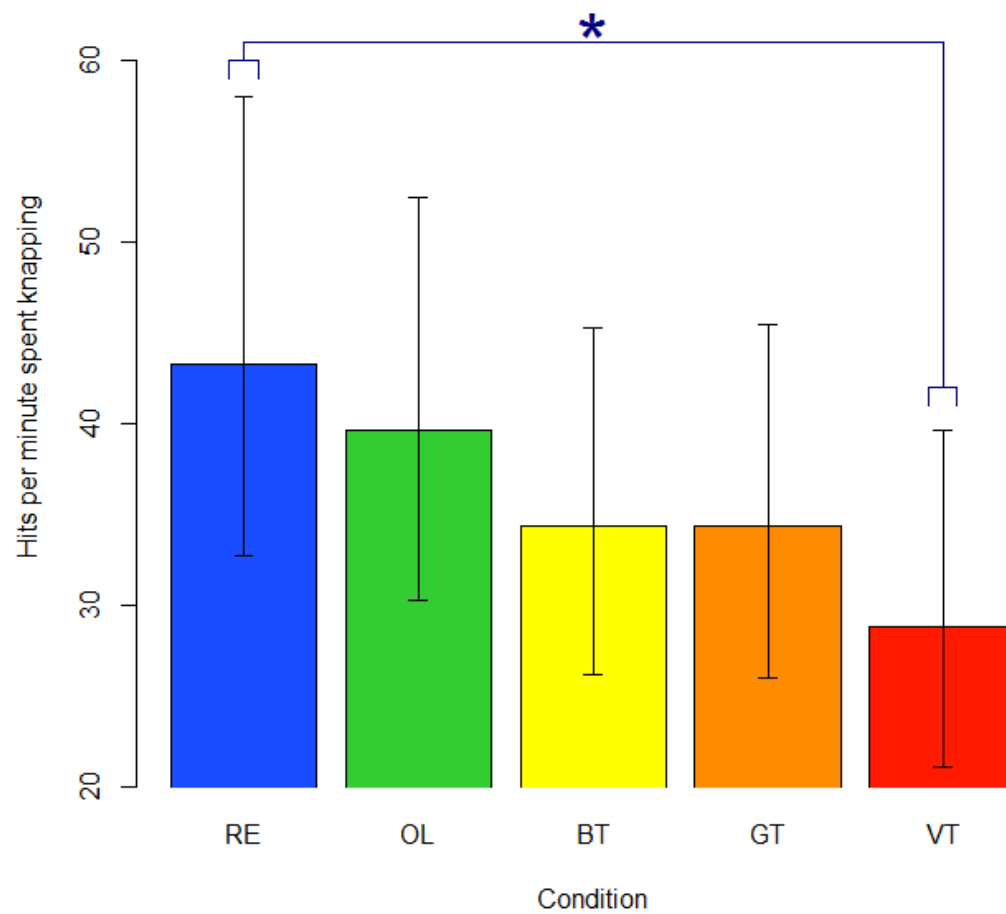


Figure 5.12

The number of times participants hit their core with their hammerstone per minute is less with Verbal Teaching than with Reverse Engineering. The brackets indicate contrasts for which there is evidence of a difference, single asterisks correspond to weak evidence, double asterisks to strong evidence. The height of each bar is the median model estimate, whilst the error bars show the 95% central credible interval.

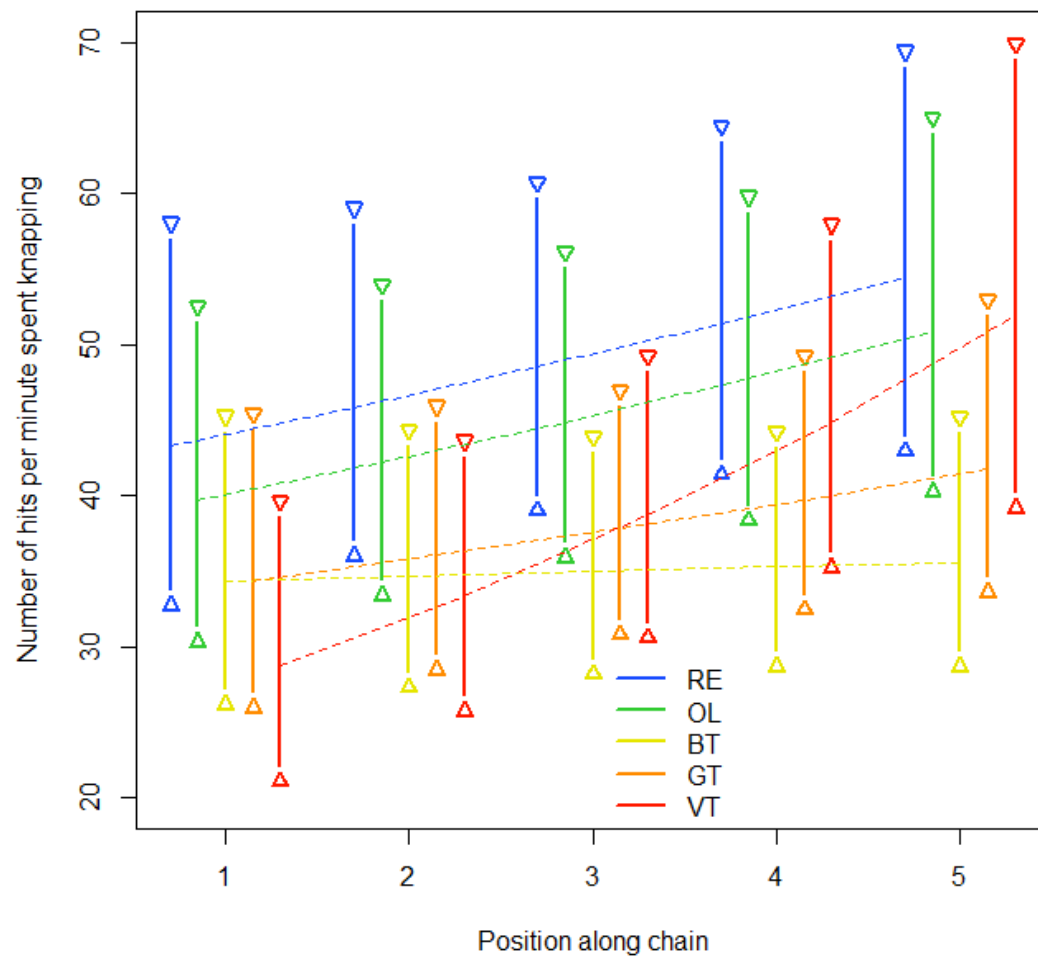


Figure 5.13

The number of times participants hit their core with their hammerstone per minute spent knapping shows a general tendency to increase along chains within conditions, but there is only strong evidence for this effect with verbal teaching (red lines). The solid lines represent the 95% central credible intervals for estimated values at the first five points in chains for each condition, whilst the dashed lines link their medians to show the trend along chains.

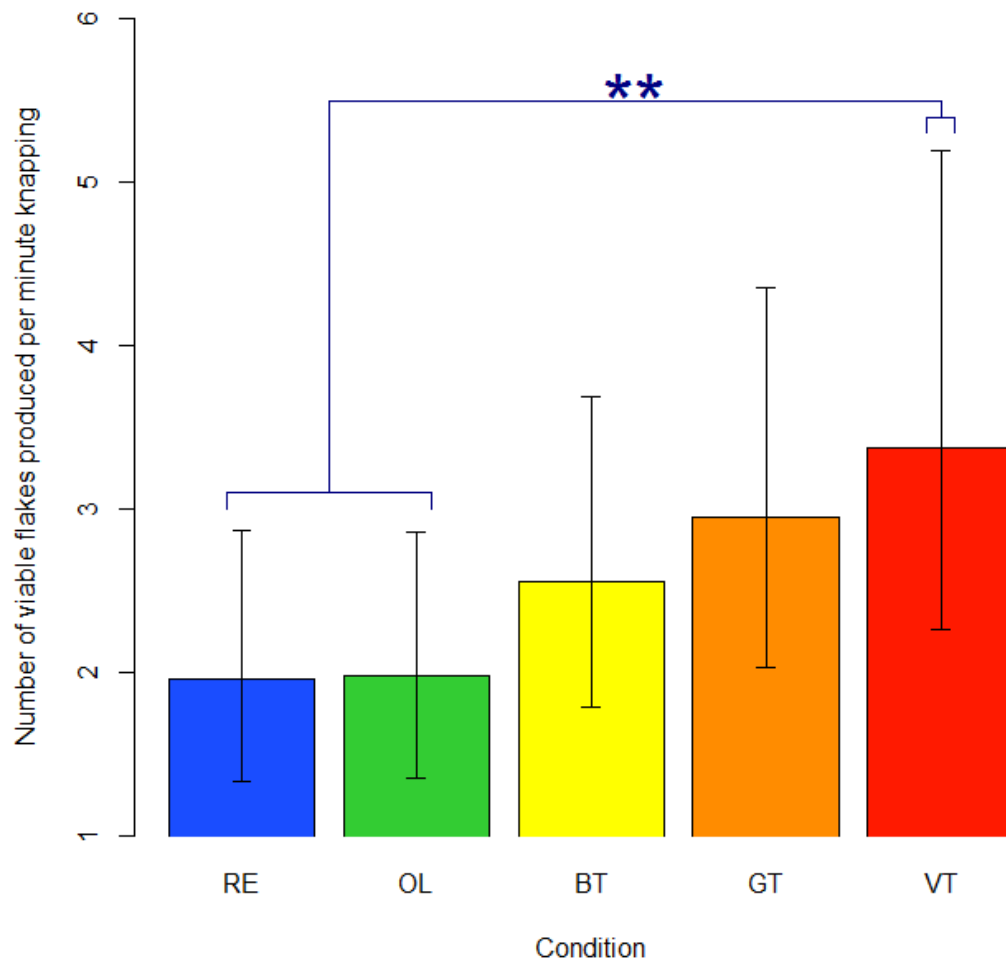


Figure 5.14

The number of viable flakes participants were able to produce per minute was greater with Verbal Teaching than with Reverse Engineering or Observational Learning. The brackets indicate contrasts for which there is evidence of a difference, single asterisks correspond to weak evidence, double asterisks to strong evidence. The height of each bar is the median model estimate, whilst the error bars show the 95% central credible interval.

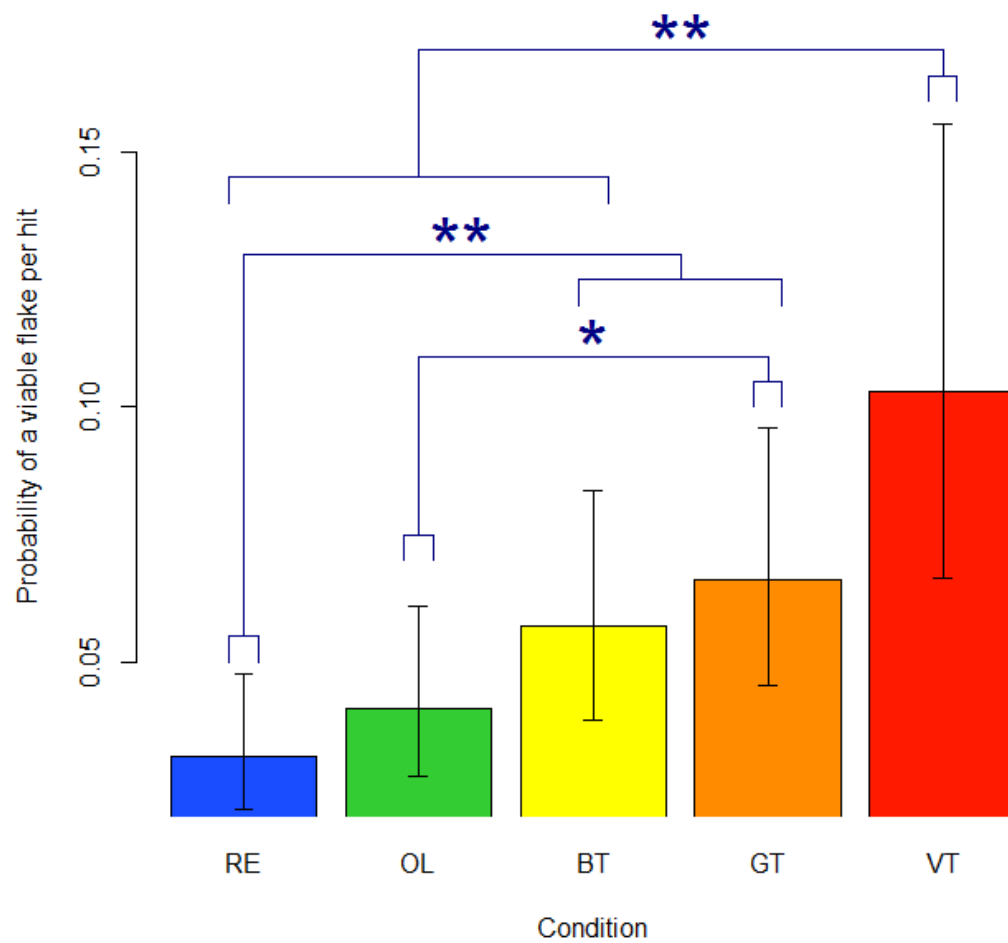


Fig 5.15

The probability that each hit produces a viable flake increases across conditions. The brackets indicate contrasts for which there is evidence of a difference, single asterisks correspond to weak evidence, double asterisks to strong evidence. The height of each bar is the median model estimate, whilst the error bars show the 95% central credible interval.



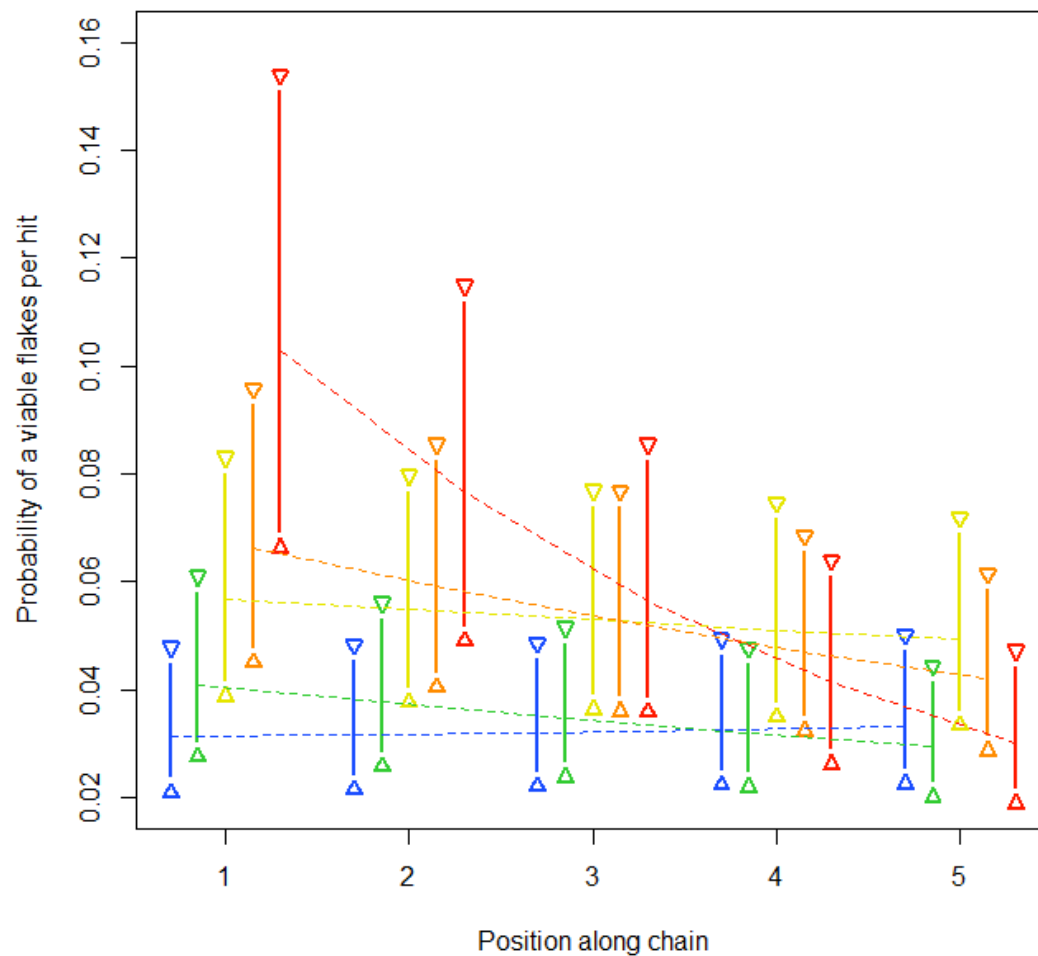


Figure 5.16

The probability of a viable flake per hit decreases along chain such that although there are differences between conditions at the start of the chains, by position 5 there are no such differences. There is evidence of this decline in all conditions other than Reverse Engineering (although the evidence is only weak in the case of Verbal Teaching). The solid lines represent the 95% central credible intervals for estimated values at the first five points in chains for each condition, whilst the dashed lines link their medians to show the trend along chains.

There is strong evidence that the **amount of time tutors spent knapping** in their tutor phase was lower with VT than BT or GT and lower with GT than BT. In all conditions that amount of time tutors spent knapping increased along chains. There was a positive effect of tutor quality on the time tutors spent knapping with BT, but a negative effect with VT and GT (for these values and all other estimated effects of tutor quality on tutor and pupil behaviour, see **Table 5.5**). There is strong evidence that the **amount of time tutors spent observing their pupil** was higher with VT than with GT or BT and higher with GT than BT. In all conditions that amount of time tutors spent observing their pupil decreased along chains. There was a negative effect of tutor quality on the amount of time tutors spent observing their pupil with BT, but a positive effect with GT and VT. There is strong evidence that the **amount of time tutors spent teaching their pupil** was higher with VT than GT or BT and higher with GT than BT. In all conditions there was a negative effect of position along the chain on the amount of time tutors spent teaching their pupil. There is a positive effect of tutor quality on the amount of time tutors spent teaching their pupil with BT and GT, but a negative effect with VT. **The amount of time tutors spent choosing between flakes** during their tutor phase was less with VT than with GT or BT and less with GT than with BT. There was a negative effect of position along the chain on the amount of time tutors spent choosing between flakes with BT, but a positive effect with GT and VT. There is a negative effect of tutor quality on the amount of time tutors spent choosing between flakes with GT, but there is not strong evidence of an effect with BT or VT. There is strong evidence that the **number of times tutors hit their core with their hammerstone** during their tutor phase was fewer with VT than with GT or BT and fewer with GT than with BT. There was an increase along chains in all conditions in the number of times tutors hit their core with their hammerstone.

There was a positive effect of tutor quality on the number of times tutors hit their core with their hammerstone with BT, but a negative effect with GT and VT. There is strong evidence that the **amount of time pupils spent knapping** in their pupil phase was less with VT than in any other condition and less with GT or BT than OL, but there is not strong evidence of a difference between BT and GT. There was a negative effect of position on the amount of time pupils spent knapping with OL, but a positive effect in the other conditions. There was a positive effect of tutor quality on the amount of time pupils spent knapping with BT and VT, but a negative effect with OL and GT. There is strong evidence that the **amount of time pupils spent observing their tutor** differs between all conditions and is greater with more complex communication. There was a positive effect of chain position on the amount of time pupils spent observing their tutor with OL, no effect with BT, and a negative effect with GT and VT. There was a positive effect of tutor quality on the amount of time pupils spent observing their tutor with OL, but a negative effect in all other conditions. There is strong evidence that the **amount of time pupils spent choosing between flakes** during their pupil phase was lower with VT than any other form of communication and lower with GT than with BT or OL, but there is not strong evidence for a difference between BT and OL. There was a negative effect of position along the chain on the amount of time pupils spent choosing between flakes with OL and BT, and a positive effect with GT, but there is not strong evidence of an effect with VT. There was a positive effect of tutor quality on the amount of time pupils spent choosing between flakes with OL and VT, but a negative effect with BT and there is not strong evidence of an effect in either direction with GT. There is strong evidence that the **number of times pupils hit their core with their hammerstone** in their pupil phase was fewer with VT than any other condition, fewer with BT than

with OL, but greater with GT than with BT. There is not strong evidence of a difference between GT and OL. Position along the chain had a positive effect on the number of times pupils hit their core with their hammerstone in all conditions, other than with BT where there is little evidence of an effect in either direction. Tutor quality had a positive effect on the number of times pupils hit their core with their hammerstone with VT and BT, but a negative effect with GT and OL.

### 5.5.7 - Language

As described in the analysis (**Section 5.4.7**), the analyses of language involved a rate function that set a baseline value, initial deviation from baseline and rate of approach from initial value to baseline. To aid interpretation of the rate parameter, a value greater than 2 is very rapid rate of approach such that ~90% of any change is achieved in the first step. A value below 0.5 corresponds to a more gentle change with ~90% of the change occurring over the first 5 steps, and lower values correspond to even gentler change, values between 2 and 0.5 correspond to intermediate rates of change. The **total number of utterances** rapidly decreased to baseline along chains (see **Fig. 5.17**). The **categories** that showed no change along chains in the probability they described an utterance are “said-by-the-tutor”, “teaching”, “feedback”, “watch-me” and “abstract”, whilst “requesting-information”, “conveying-uncertainty” and “incorrect” increased and “confirmation-of-understanding”, “this/that” and “correct” (see **Fig. 5.18**) decreased. The **topics** that showed no change along chains in the probability that they were covered in an utterance are “platform-edge”, “how-to-hit”, “force”, “how-to-hold”, “hammerstones”, “cortex”, “cutting-edge” and “flake-size”,

whilst “knapping”, “knapping-site”, “platform-angle”, “ridge” and “choosing-flakes” decreased (for these values see **Table 5.7**). No topics increased along chains in the probability of being covered by an utterance. The various topics showed a range of accuracies (see **Table 5.6**). There is evidence that **the probability information is correct** is lower for utterances covering the knapping site than for utterances covering the platform edge, cortex or flake cutting edge. There is also evidence that the probability information about the ridge is correct is greater than for all other topics. There is no evidence for any other differences.

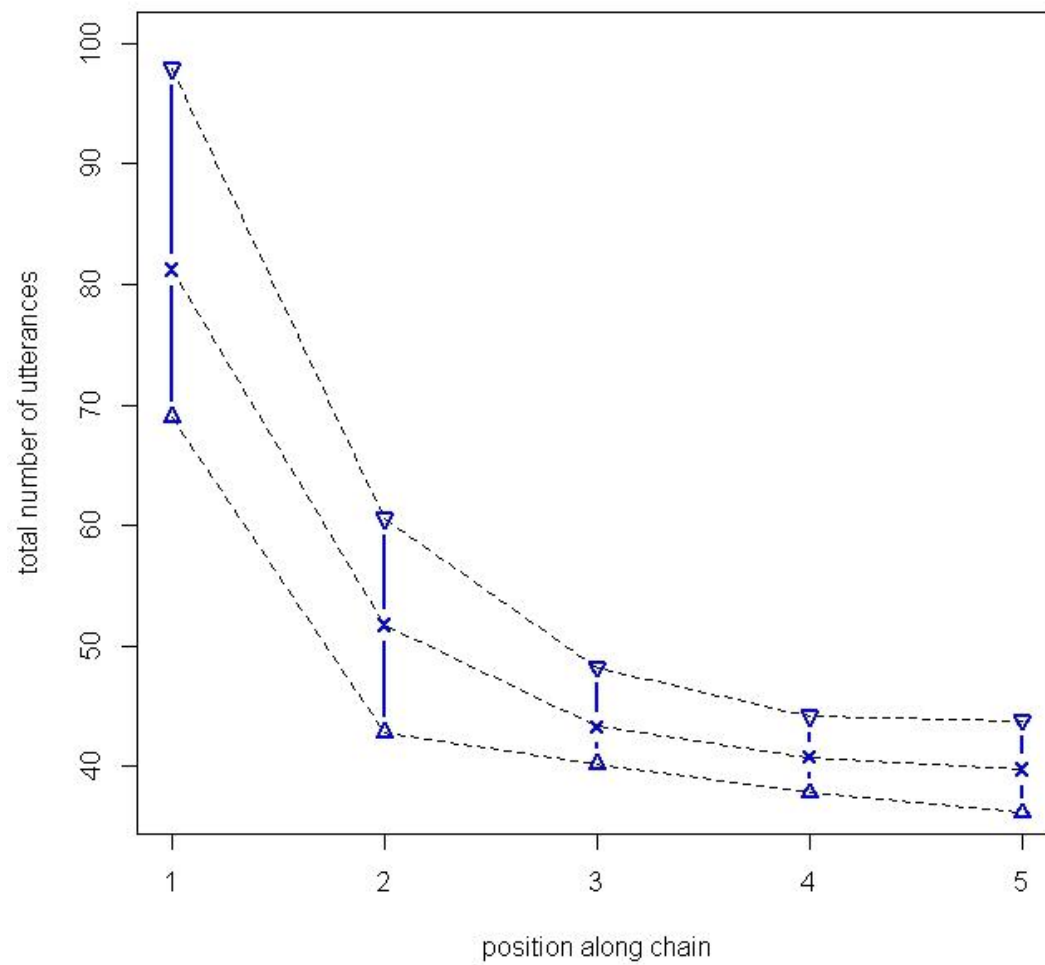


Fig 5.17: The total number of utterances said by tutors and pupils decreases along chains. The values shown are median model estimates and their associated 95% central credible interval.

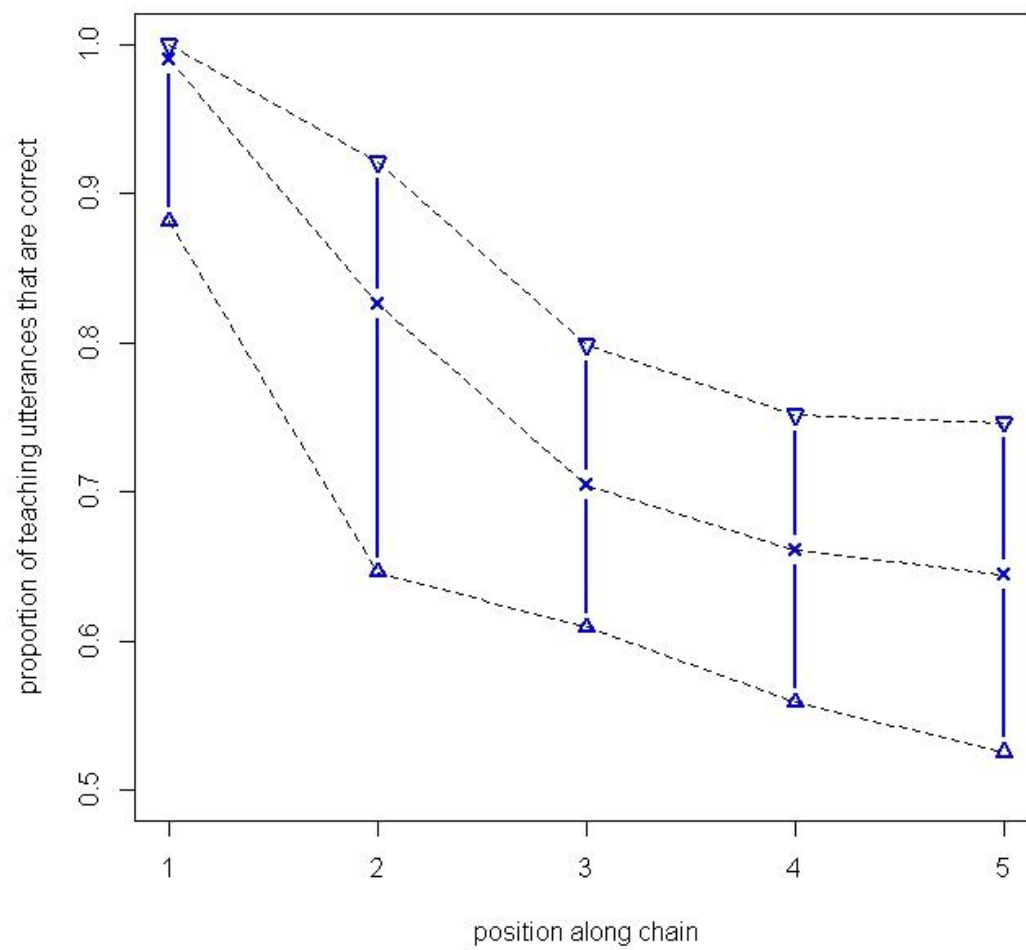


Fig 5.18

The probability that a teaching utterance contains correct information decreases along chains. The values shown are median model estimates and their associated 95% central credible interval.

Table 5.1

Estimated values for parameters at the first position in the chain for different conditions.  
Quoted values are medians and 95% central credible intervals.

Variable		Condition				
		RE	OL	BT	GT	VT
Flake mass (g)	All	12.6, [9.8, 16.2]	13.1, [10.2, 6.7]	19.8, [15.5, 5.6]	16.8, [13.1, 1.5]	14.7, [11.4, 9.0]
	Viable	14.8, [11.4, 9.3]	14.2, [10.9, 8.5]	20.9, [16.1, 7.3]	18.7, [14.5, 4.2]	15.8, [12.2, 0.4]
	Non-viable	9.4, [6.9, 13.0]	10.9, [8.1, 14.6]	14.2, [10.4, 9.8]	10.9, [7.9, 15.0]	10.4, [7.5, 14.0]
	Selected	17.5, [12.0, 5.6]	19.5, [13.3, 7.8]	25.9, [18.1, 7.0]	29.7, [20.6, 3.4]	15.8, [10.7, 2.9]
	Non-selected	7.2, [5.2, 10.0]	7.9, [5.7, 10.7]	9.5, [7.0, 13.3]	7.6, [5.5, 10.6]	8.6, [6.3, 12.1]
Flake Diameter (cm)	All	3.9, [3.6, 4.3]	4.0, [3.7, 4.3]	4.7, [4.3, 5.0]	4.4, [4.1, 4.8]	4.2, [3.9, 4.6]
	Viable	4.3, [3.9, 4.6]	4.3, [4.0, 4.7]	4.8, [4.5, 5.2]	4.7, [4.3, 5.0]	4.3, [4.0, 4.7]
	Non-viable	3.5, [3.2, 3.8]	3.6, [3.3, 3.9]	4.1, [3.7, 4.5]	3.7, [3.4, 4.1]	3.7, [3.4, 4.0]
	Selected	4.6, [4.1, 5.2]	4.8, [4.3, 5.4]	5.3, [4.8, 5.9]	5.6, [5.0, 6.3]	4.5, [4.0, 5.0]
	Non-selected	3.4, [3.1, 3.8]	3.5, [3.2, 3.8]	3.8, [3.5, 4.2]	3.6, [3.3, 4.0]	3.7, [3.4, 4.1]
Flake cutting edge (cm)		3.0, [2.7, 3.4]	3.2, [2.8, 3.6]	3.2, [2.8, 3.6]	3.5, [3.1, 4.0]	3.4, [3.0, 3.8]
Flake quality		0.76, [0.65,0.89]	0.85, [0.73,0.98]	0.81, [0.70,0.94]	0.88, [0.76,0.02]	0.84, [0.72,0.97]
Number of flakes	All	28.0, [21.9,36.0]	31.7, [24.9,40.5]	27.9, [21.8,35.3]	30.1, [23.5,38.4]	34.3, [26.9,43.8]
	Viable	15.8, [12.1, 0.5]	18.3, [14.1, 3.6]	19.6, [15.1, 5.4]	21.7, [16.8, 8.3]	25.2, [19.4,33.0]
	Non-viable	12.0, [9.1, 15.9]	13.1, [10.0,17.1]	8.1, [6.1, 10.9]	8.6, [6.5, 11.3]	9.6, [7.2, 12.7]
	Selected	12.5, [9.4, 16.4]	13.3, [10.1,17.4]	16.3, [12.5,21.1]	14.8, [11.3,19.4]	23.0, [17.5,30.4]
	Non-selected	14.7, [11.3,19.3]	17.6, [13.6,23.0]	11.3, [8.6, 14.7]	14.6, [11.3,19.0]	13.1, [10.1,17.1]
Proportion of flakes	Viable	0.55, [0.48,0.62]	0.58, [0.52,0.64]	0.72, [0.66,0.77]	0.72, [0.67,0.77]	0.73, [0.68,0.78]
	selected	0.46, [0.39,0.53]	0.45, [0.38,0.51]	0.62, [0.55,0.68]	0.48, [0.42,0.55]	0.61, [0.54,0.67]
Total cutting edge (cm)		52.6, [37.3,72.3.]	61.3, [43.5,84.0]	62.3, [46.2,83.2]	81.2, [59.7,109.5]	98.1, [72.0,133.3]
Total flake mass (g)		40.6, [28.2,55.8]	45.1, [31.1,62.2]	57.1, [41.2,76.3]	59.7, [42.8,80.9]	59.3, [42.3,79.9]
Total quality		13.0, [9.2, 17.9]	15.7, [11.1,21.4]	15.4, [11.1,20.7]	19.8, [14.6, 26.7]	23.6, [17.0, 31.9]



Proportion of core remaining		0.56, [0.46,0.65]	0.54, [0.44,0.63]	0.47, [0.37,0.57]	0.49, [0.38,0.57]	0.41, [0.29,0.52]
Hits per minute knapping		43.2, [32.7,57.5]	39.7, [30.1,52.5]	34.5, [26.1,45.2]	34.3, [26.0,45.5]	28.8, [20.9,39.3]
Flakes per minute	All	3.28, [2.31,4.62]	3.13, [2.21,4.36]	3.56, [2.56,5.00]	4.04, [2.87,5.77]	4.52, [3.15,6.69]
	viable	1.96, [1.33,2.87]	1.98, [1.35,2.85]	2.55, [1.78,3.69]	2.95, [2.03,4.36]	3.37, [2.26,5.19]
Probability of a viable flake per hit		0.03, [0.02,0.05]	0.04, [0.03,0.06]	0.06, [0.04,0.08]	0.07, [0.05,0.10]	0.10, [0.07,0.16]

Table 5.2

Values for contrasts between estimates for variables between conditions. Values are estimated as the expected value for the first condition/topic minus the expected value for the second condition/topic. Quoted values are median estimates with 95% central credible intervals. Due to the large number of analyses and variables considered I only present values for contrasts where there is evidence of a difference (where the numbers quoted are in *italics* this means that the evidence is only weak, as opposed to strong). In some cases the 1st condition/topic and 2nd condition/topic columns contain two values, in this case the value given is for a contrast between contrasts. The contrast between the pair of conditions given in each condition column was first estimated and then the difference between these two contrasts estimated.

Variable	First condition		Second condition		Contrast
Flake mass	BT		RE		7.13, [1.66, 13.55]
			OL		6.69, [1.17, 13.2]
Viable flake mass	BT		OL		6.60, [0.36, 13.84]
Selected flake mass	GT		RE		12.14, [0.23, 27.14]
			VT		13.69, [2.45, 28.65]
Flake diameter	BT		RE		0.71, [0.27, 1.16]
			OL		0.64, [0.21, 1.1]
	GT		RE		0.48, [0.03, 0.91]
Viable flake diameter	BT		RE		0.56, [0.06, 1.08]
			OL		0.50, [0.01, 1.00]
Non-viable flake diameter	BT		RE		0.58, [0.12, 1.09]
			OL		0.51, [0.03, 1.00]
Selected flake diameter	GT		RE		1.01, [0.20, 1.84]
			VT		1.17, [0.38, 1.97]
	BT		VT		0.84, [0.09, 1.58]
Number of viable flakes	VT		RE		9.4, [2.1, 18.1]
			OL		6.9, [-0.8, 18.1]
	GT		RE		6.0, [-0.7, 13.5]
Proportion of flakes that are viable	VT		RE		0.18, [0.12, 0.25]
			OL		0.15, [0.09, 0.21]
	GT		RE		0.17, [0.11, 0.24]
			OL		0.14, [0.08, 0.20]
	BT		RE		0.17, [0.11, 0.23]
			OL		0.14, [0.08, 0.20]
	BT	OL	VT	GT	0.57, [0.20, 0.95]
			GT	BT	0.60, [0.13, 1.08]
			OL	RE	0.49, [0.05, 0.94]
	Number of non-viable flakes	GT		RE	
OL				-4.5, [-9.1, -1.0]	
BT		RE		-3.8, [-8.3, 0.1]	
		OL		-4.9, [-9.6, -0.8]	
Number of selected flakes	VT		GT		8.1, [1.2, 16.3]
			BT		6.7, [-0.5, 14.8]
			OL		9.6, [2.7, 17.6]
			RE		10.5, [3.6, 18.5]
Proportion of	VT		GT		0.12, [0.06, 0.18]

flakes that were selected	BT	OL	0.16, [0.10, 0.22]
		RE	0.15, [0.08, 0.22]
		GT	0.13, [0.07, 0.20]
		OL	0.17, [0.11, 0.23]
		RE	0.16, [0.09, 0.22]
Number of non-selected flakes	BT	OL	-6.3, [-12.6, -1.0]
Total quality	VT	BT	8.2, [-0.1, 17.4]
		OL	7.9, [-1.1, 17.5]
		RE	10.6, [2.2, 20.0]
	GT	RE	6.7, [-0.4, 14.7]
Total cutting edge	VT	BT	36.0, [2.7, 72.9]
		OL	36.6, [2.9, 76.4]
		RE	45.7, [12.0, 85.4]
	GT	RE	28.4, [-0.3, 61.3]
Total mass	RE	VT	-18.6, [-41.6, 2.0]
		GT	-18.9, [-40.8, 0.29]
		BT	-16.2, [-36.1, 1.9]
Proportion of core remaining	VT	RE	-0.15, [-0.31, -0.00]
		OL	-0.13, [-0.29, 0.01]
Hits per minute knapping	VT	RE	-14.3, [-30.8, -0.11]
Viable flakes per minute knapping	VT	RE	1.39, [0.03, 3.35]
		OL	1.37, [0.03, 3.34]
Probability of a viable flake with each hit	VT	BT	0.05, [0.00, 0.10]
		OL	0.06, [0.02, 0.12]
		RE	0.07, [0.03, 0.12]
	GT	OL	0.02, [-0.00, 0.06]
		RE	0.03, [0.01, 0.07]
	BT	RE	0.03, [0.00, 0.05]
Topic Accuracy	Ridge	Knapping site	0.44, [0.22, 0.66]
		Platform edge	0.07, [0.01, 0.20]
		Platform angle	0.28, [0.06, 0.63]
		How to hit	0.20, [0.06, 0.42]
		How to hold	0.16, [0.03, 0.47]
		Hammerstones	0.27, [0.09, 0.52]
		Cortex	0.06, [-0.00, 0.23]
		Flake size	0.31, [0.11, 0.60]
		Cutting edge	0.08, [0.02, 0.19]
		Force	0.61, [0.39, 0.79]
	Cortex	Knapping site	0.37, [0.09, 0.62]
		Force	0.54, [0.28, 0.74]
	Platform edge	Knapping site	0.37, [0.07, 0.62]
		Flake size	0.24, [0.00, 0.53]
		Force	0.53, [0.30, 0.73]
		Hammerstones	0.19, [-0.01, 0.40]
	Cutting edge	Knapping site	0.35, [0.12, 0.58]
		Hammerstones	0.18, [-0.01, 0.44]

		Flake size	0.23, [-0.02, 0.54]
		Force	0.52, [0.29, 0.72]
	Force	How to hit	-0.41, [-0.66, -0.08]
		How to hold	-0.43, [-0.68, -0.08]
		Hammerstones	-0.33, [-0.60, -0.02]
		Flake size	-0.29, [-0.56, 0.04]

Table 5.3

Estimated values for effects of position along the chain on different variables and for different conditions. Quoted values are medians and 95% central credible intervals.

Variable		Condition				
		RE	OL	BT	GT	VT
Flake mass (g)	All	-0.01, [-0.07,0.06]	0.034, [-0.03,0.10]	-0.04, [-0.11,0.02]	-0.02, [-0.09,0.05]	-0.01, [-0.08,0.05]
	Viable	0.00, [-0.07, 0.07]	0.04, [-0.02, 0.11]	-0.03, [-0.09,0.04]	-0.02, [-0.09, 0.05]	0.00, [-0.06, 0.07]
	Non-viable	-0.02, [-0.11, 0.06]	0.01, [-0.07, 0.09]	-0.06, [-0.14,0.01]	0.01, [-0.07, 0.10]	-0.01, [-0.09, 0.07]
	Selected	-0.03, [-0.13, 0.07]	0.09, [-0.00, 0.19]	0.03, [-0.06,0.12]	-0.06, [-0.16, 0.05]	0.08, [-0.02, 0.17]
	Non-selected	0.02, [-0.07, 0.11]	-0.05, [-0.13, 0.04]	-0.04, [-0.13,0.03]	0.01, [-0.08, 0.10]	-0.06, [-0.15, 0.02]
Flake Diameter (cm)	All	-0.00, [-0.02, 0.02]	0.01, [-0.01, 0.03]	-0.01, [-0.03,0.00]	-0.01 [-0.03,0.01]	-0.01, [-0.03, 0.01]
	Viable	-0.00, [-0.02, 0.02]	0.01, [-0.01, 0.03]	-0.01, [-0.03,0.01]	-0.01, [-0.03, 0.01]	-0.00, [-0.02, 0.02]
	Non-viable	-0.01, [-0.03, 0.02]	0.01, [-0.01, 0.03]	-0.02, [-0.04,0.00]	0.00, [-0.02, 0.02]	-0.01, [-0.04, 0.01]
	Selected	-0.01, [-0.04, 0.02]	0.02, [-0.01, 0.05]	0.01, [-0.02,0.04]	-0.03, [-0.05, 0.01]	0.02, [-0.01, 0.05]
	Non-selected	0.01, [-0.02, 0.03]	-0.01, [-0.03, 0.02]	-0.01, [-0.04,0.01]	-0.00, [-0.03, 0.02]	-0.02, [-0.05, 0.00]
Flake cutting edge (cm)		0.03, [-0.01, 0.06]	0.00, [-0.03, 0.03]	0.00, [-0.03,0.03]	-0.02, [-0.05, 0.02]	0.01, [-0.02, 0.04]
Flake quality		0.03, [-0.02, 0.07]	-0.00, [-0.04, 0.03]	0.00, [-0.03,0.04]	-0.01, [-0.05, 0.03]	0.01, [-0.03, 0.04]
Number of flakes	All	0.05, [0.03, 0.08]	-0.02, [-0.05, 0.01]	0.03, [0.00, 0.05]	0.03, [0.00, 0.06]	-0.04, [-0.06,-0.01]
	Viable	0.07, [0.03, 0.10]	0.00, [-0.03, 0.04]	0.00, [-0.03,0.03]	-0.01, [-0.04, 0.02]	-0.07, [-0.10, -0.04]
	Non-viable	0.04, [-0.00, 0.08]	-0.05, [-0.09,-0.01]	0.07, [0.03,0.11]	0.09, [0.05, 0.14]	0.02, [-0.02,0.06]
	Selected	0.05, [0.01,0.09]	0.02, [-0.02,0.06]	-0.03, [-0.06,0.01]	-0.01, [-0.05,0.03]	-0.11, [-0.14,-0.07]
	Non-selected	0.06, [0.02,0.10]	-0.05, [-0.08,-0.01]	0.08, [0.04,0.11]	0.07, [0.03,0.11]	0.02, [-0.01,0.05]
Proportion of	Viable	0.03, [-0.01,0.08]	0.03, [-0.01,0.08]	-0.06, [-0.10,-0.01]	-0.11, [-0.45,-.06]	-0.08, [-0.13,-0.03]
	Selected	0.02, [-0.03,0.06]	0.02, [-0.02,0.07]	-0.12, [-0.16,-0.07]	-0.03, [-0.08,0.02]	-0.15, [-0.19,-0.10]
Total cutting edge (cm)		0.06, [-0.01,0.14]	-0.02, [-0.11,0.07]	0.02, [-0.05,0.08]	-0.04, [-0.12,0.04]	-0.06, [-0.13,0.01]
Total flake mass (g)		0.01, [-0.08,0.08]	0.01, [-0.08,0.09]	-0.01, [-0.08,0.05]	0.00, [-0.08,0.08]	-0.01, [-0.08,0.06]
Total quality		0.06,	-0.02,	0.01,	-0.04,	-0.07,

		[-0.02,0.14]	[-0.12,0.06]	[-0.05,0.07]	[-0.12,0.04]	[-0.14,0.01]
Proportion of core remaining		-0.02, [-0.13,0.08]	-0.06, [-0.16,0.04]	0.00, [-0.09,0.09]	-0.09, [-0.20,0.01]	-0.04, [-0.16,0.08]
Hits per minute knapping		0.06, [-0.01,0.13]	0.06, [-0.01,0.13]	0.01, [-0.05,0.07]	0.05, [-0.02,0.12]	0.15, [0.06,0.24]
Flakes per minute	All	0.02, [-0.07,0.11]	0.03, [-0.05,0.12]	-0.00, [-0.08,0.08]	0.00, [-0.09,0.09]	-0.09, [-0.21,0.02]
	viable	0.02, [-0.08,0.12]	0.02, [-0.07,0.12]	-0.02, [-0.11,0.07]	-0.03, [-0.13,0.07]	-0.12, [-0.25,0.00]
Probability of a viable flake per hit		0.01, [-0.02,0.05]	-0.08, [-0.12,-0.05]	-0.04, [-0.08,0.00]	-0.12, [-0.16,-.08]	-0.33, [-0.38,-0.28]

Table 5.4

Estimated values for effects of core mass on different variables. Quoted values are medians and 95% central credible intervals.

Variable		Effect of core mass
Flake mass (g)	All	0.028, [-0.054, 0.107]
	Viable	0.033, [-0.056, 0.116]
	Non-viable	0.033, [-0.069, 0.133]
	Selected	0.143, [0.020, 0.267]
	Non-selected	0.009, [-0.108, 0.119]
Flake Diameter (cm)	All	0.011 [-0.013, 0.034]
	Viable	0.014, [-0.010, 0.040]
	Non-viable	0.011, [-0.015, 0.039]
	Selected	0.047, [0.012, 0.083]
	Non-selected	0.006, [-0.025, 0.036]
Flake cutting edge (cm)		-0.019, [-0.058, 0.020]
Flake quality		-0.015 [-0.062, 0.034]
Number of flakes	All	0.13, [0.09, 0.17]
	Viable	0.13, [0.08, 0.17]
	Non-viable	0.11, [0.04, 0.17]
	Selected	-0.03, [-0.08, 0.02]
	Non-selected	0.26, [0.21, 0.31]
Total cutting edge (cm)		0.04, [-0.06, 0.15]
Total flake mass (g)		0.09, [-0.00, 0.18]
Total quality		0.05, [-0.05, 0.16]
Proportion of core remaining		-1.82, [-3.42, -0.60]

Table 5.5

Estimated values for the effect of the number of viable flakes a tutor managed to produce during their test phase on tutor and pupil behaviour in the different conditions. The quoted values are median estimates and their 95% central credible intervals. The number of flakes a tutor managed to produce was used as a measure of tutor quality.

Variable		Condition				
		RE	OL	BT	GT	VT
Time spent knapping	Tutor	na	na	0.0075, [0.0048, 0.0101]	-0.0135, [-0.0181, -0.0090]	-0.0695, [-0.0787, -0.0606]
	Pupil	na	-0.0105, [-0.0131, -0.0081]	0.0083, [0.0000, 0.0105]	-0.0028, [-0.0054, -0.0002]	0.0086, [0.0062, 0.0109]
Time spent observing	Tutor	na	na	-0.0304, [-0.0360, -0.0247]	0.0043, [0.0008, 0.0078]	0.0143, [0.0116, 0.0170]
	Pupil	na	0.0064, [0.0024, 0.0103]	-0.0132, [-0.0164, -0.0099]	-0.0043, [-0.0077, -0.0009]	-0.0152, [-0.0181, -0.0123]
Time spent teaching	Tutor	na	na	0.0050, [0.0002, 0.0097]	0.0064, [0.0239, 0.0102]	(-0.0068, [-0.0097, -0.0039]
Time spent choosing	Tutor	na	na	0.0036, [-0.0008, 0.0080]	-0.0127, [-0.0193, -0.0060]	0.0040, [-0.0131, 0.0201]
	Pupil	na	0.0272, [0.0225, 0.0319]	-0.0048, [-0.0093, -0.0003]	-0.0060, [-0.0120, 0.0000]	0.0162, [0.0103, 0.0218]
Number of hits	Tutor	na	na	0.019, [0.016, 0.022]	-0.024, [-0.029, -0.018]	-0.058, [-0.066, -0.051]
	Pupil	na	-0.0213, [-0.0245, -0.0181]	0.0126, [0.0095, 0.0157]	-0.0075, [-0.0108, -0.0043]	0.0064, [0.0032, 0.0094]



Table 5.6

Estimated values for average topic accuracy (i.e., the probability that an utterance that contains information on that topic contains factually correct information). Quoted values are median estimates with their 95% central credible intervals.

topic	Accuracy
ridge	1.00, [0.95, 1.00]
cortex	0.91, [0.71, 0.99]
platform edge	0.89, [0.73, 0.96]
flake cutting edge	0.88, [0.75, 0.95]
how to hold	0.77, [0.45, 0.95]
how to hit	0.71, [0.48, 0.88]
hammerstones	0.67, [0.44, 0.86]
flake size	0.67, [0.39, 0.88]
platform angle	0.62, [0.30, 0.89]
knapping site	0.54, [0.32, 0.75]

Table 5.7

Estimated values for rate and extent of change for variables concerning verbal interactions along chains. A negative value for change corresponds to a decrease along the chain. Quoted values are medians and 95% central credible intervals.

Variable/Category/Topic	Rate of change along chains	Change along chains
Total Utterances	1.2, [0.63, 14.0]	-42.2, [-29.3, -58.9]
Proportion of teaching utterances correct	1.4, [0.56, 45.8]	-4.0, [-1.4, -6.9]
Platform angle teaching accuracy	3.99, [0.0, 128.1]	-0.75, [3.21, -1.91]
Ridge teaching accuracy	0.42, [0.1766, 1.10]	-3.69, [-1.95, -6.75]
Platform edge teaching accuracy	0.00, [0.0, 0.09]	1.18, [4.78, -4.12]
Force required teaching accuracy	0.00, [0.0, 0.03]	0.53, [4.73, -3.489]
Said by the tutor	0.00, [0.0, 0.00]	-0.76, [-3.57, 5.19]
Teaching	0.00, [0.0, 0.01]	-0.28, [-5.76, 3.87]
Feedback	0.00, [0.0, 0.06]	-0.28, [-3.90, 3.25]
Confirmation of understanding	13.33, [1.89, 163.5]	-0.88, [-1.77, -0.09]
Watch this	0.00, [0.0, 0.30]	2.35, [-2.99, 6.47]
This/that	0.40, [0.00, 91.57]	-0.56, [-3.35, 3.56]
Requesting Information	10.92, [0.86, 149.5]	0.96, [-0.04, 2.23]
Conveying uncertainty	7.18, [1.63, 159.0]	3.88, [1.95, 6.69]
Abstract	0.00, [0.0, 0.00]	-0.52, [-4.40, 3.15]
Correct	4.03, [1.38, 6.90]	-4.03, [-6.90, -1.38]
Incorrect	2.36, [0.83, 98.85]	4.00, [-1.33, 7.39]
Knapping	0.11, [0.00, 111.0]	-0.74, [-4.07, 2.08]
Knapping site	0.09, [0.02, 7.82]	-2.31, [-5.65, -0.54]
Platform edge	0.00, [0.0, 0.09]	1.18, [-4.13, 4.78]
Platform angle	3.99, [0.0, 128.1]	-0.75, [-1.91, 3.21]
Ridge	0.42, [0.18, 1.10]	-3.69, [-6.75, -1.95]
force	0.00, [0.0, 0.03]	0.53, [-3.49, 4.37]
How to hit	0.00, [0.0, 0.0]	1.01, [-4.01, 5.52]
Hot to hold	0.00, [0.0, 0.00]	0.68, [-3.93, 4.68]
Hammerstones	0.00, [0.0, 1.51]	1.72, [-1.81, 6.25]
Cortex	0.00, [0.0, 0.65]	1.79, [-2.16, 6.72]
Choosing flakes	9.97, [0.00, 161.8]	0.82, [-1.73, 3.73]
Size of flakes	0.00, [0.0, 0.00]	2.01, [-1.94, 6.15]
Cutting edge of flakes	0.00, [0.0, 0.09]	1.09, [-2.64, 6.09]

## **5.6 - Discussion**

I carried out a transmission chain experiment to investigate the social transmission of the ability to produce Oldowan stone tools. Due to the number of results to discuss I shall first discuss each set of results using the same categories as in previous sections before drawing out the overall findings of this work.

### **5.6.1 - Aggregate flakes**

A central finding of this experiment is that more complex communication allowed participants to produce a greater number of viable flakes; in that evidence for a difference between several conditions was found and in every case the difference favoured the more complex condition. It should be noted, however, that there was not always evidence of a difference between all pairs of conditions. There was an increase across conditions in the number of all flakes, viable flakes, and selected flakes produced, and a decrease in the number of non-viable flakes produced and no evidence for a trend in the number of non-selected flakes produced. The effect was strongest in the case of the number of viable flakes produced, which in turn likely drove the parallel increase seen in the number of flakes that were selected by the participant. A similar trend is seen if we consider the proportion of flakes that were viable, and a similar, though weaker, trend in the case of the proportion of flakes that are selected. The proportion of flakes that are non-viable or not-selected must,

therefore, decrease, being the inverse of the proportion of flakes that are viable and selected respectively.

Furthermore, there is also clear evidence of an increase in the total quality and cutting edge of flakes that participants were able to produce - a result that can be expected given the increases described in the previous paragraph. Although there is a positive trend in the case of the total mass of flakes produced, without strong evidence for differences between conditions it is hard to draw a strong conclusion in this case. However, given the tasks such tools were used for (e.g., butchery) the importance of mass produced in stone knapping is plausibly secondary to the cutting edge and quality and so and overall positive effect of the complexity of communication on the ability to get the most out of a single piece of flint is clear.

This is a particularly strong finding as the production of a large number of viable flakes is what determined the payment of participants and so it is likely that all participants were motivated to produce as many as they could. Thus, I can rule out that the inferior performance seen in the conditions with more restricted communication reflects poor participant motivation.

### 5.6.2 - Core remaining

There is a negative trend in the proportion of the core remaining after knapping with more complex communication across conditions. This is intuitive given that with more complex communication participants were able to produce a greater number of flakes than with less complex communication. However, it is of note that any material was left behind at all given that participants were given a large amount of time in which to knap flakes and that they did not typically use all of this time. This implies, that participants abandoned their core, not because they ran out of time, but because they felt they could proceed no further. Observation of the remaining cores makes it clear why this was the case; they are typically covered with many impact marks caused by the participants striking the core, but failing to detach a flake. This is no surprise given the findings of the video analyses that found subjects typically hit the core between 10 and 30 times (depending on condition) in order to produce each viable flake. Each unsuccessful hit will damage the core preventing further knapping from that point and, without the ability to repair such damage, over time this will leave the core with no suitable knapping sites remaining on the core, forcing participants to give up. With sufficient knowledge it is possible to repair damage done to the core, as is evidenced even in the early archaeological record (Delagnes and Roche, 2005; Gowlett, 1979), although this was clearly beyond the ability of our participants and given the low success rate of each hit it is likely they would damage cores at a greater rate than it is possible to repair. Despite this, it is clear that more complex communication helped participants in that participants from chains with more complex communication were able to knap more of their core, presumably

because they did not damage it as rapidly as participants in the chains with less complex communication.

### **5.6.3 - Videos**

Similar patterns to those seen in the aggregate flake measures and the proportion of core remaining are seen in the results of analyses of the video data. With more complex communication participants were able to knap more efficiently both with regards to time, and per hit. In particular the probability that each hit produces a viable flake shows a dramatic increase across conditions, with participants in the VT condition almost 4 times as likely to produce a viable flake with each hit than those in the RE condition.

The videos also have the potential to shed light on what it was that the different conditions permitted that allowed better transmission of information to occur. For example, with increasing communication tutors taught and observed their pupil more, but knapped and chose between their own flakes less. Correspondingly, pupils spent more time observing their tutor and less time interacting with their own core and hammerstone. Presumably this is because with less complex communication, tutors were unable to convey much information and so spent less time trying to do so, or had to resort to other, less effective, means to do so. It is also likely that as information was lost along chains in these conditions tutors further down the chain had less information to impart and so there was even less point in teaching that at the start of

chains. The decrease in behaviours directed towards their own knapping materials with more complex communication shows how tutors and pupils in the less complex conditions, unable to teach or learn socially effectively, focussed instead on their own material. Indeed, from observation of the experiment, the pupil phase often lost the tutor-pupil dynamic and instead became a collaborative exercise in which each participant investigated their own materials themselves, but tried to share information and ideas as they occurred to them where possible.

#### **5.6.4 - Language**

The use of language in the verbal teaching condition can reveal how language was of benefit to participants, which in turn has the potential to reveal what the benefit of language would have been to early hominin knappers. The most striking finding is the rapid decrease in the number of things that were said along chains. This mirrors a drop in other metrics of performance along chains, such as the number of viable flakes produced. Several of the categories and topics also showed a drop in how often they were talked about along chains, most notably the proportion of utterances that were correct decreased, highlighting the ability of language to accumulate and transmit incorrect information between individuals. Many of the topics that decreased in their frequency of use were related to where to strike the core with the hammerstone (e.g., knapping-site, platform-angle, ridge). It is tempting to suggest that the loss of such information could be behind the decrease in performance as such information is very important to successful knapping. Supporting this, knapping-site, another important topic, also accumulated errors more rapidly than other topics. Although the presence

of a ridge was a topic that accumulated very few errors, this seems to be because it disappeared so quickly that there were insufficient transmission opportunities in which errors could accumulate. Despite this, as only one condition involved language and without an assay of participant knapping knowledge, it is difficult to verify what it was that language enabled the transmission of.

### **5.6.5 - Flake identification**

Across conditions there are some trends in the ability of participants to identify flakes successfully. The probability that a selected flake was viable increased across conditions, suggesting an increase in the ability to identify flakes with more complex communication. However the probability that a non-selected flake is non-viable decreased with more complex communication. Given that there was no evidence of a general trend across conditions in the probability a viable flake was selected, any improvement in ability to identify flakes seems subtle. Indeed, the increase in the probability that a selected flake was viable (or the decrease in the probability that a non-selected flake was non-viable) across conditions may just reflect the fact that a greater proportion of the flakes produced by participants were viable as the complexity of communication increased. Even if participants were selecting flakes at random, with more viable flakes available to participants (both in number and proportion) in the conditions with more complex communication, a greater proportion of selected flakes would be viable than in conditions with less complex communication and fewer viable flakes produced. I do not mean to suggest participants were genuinely selecting flakes at random; however, it is odd that with



more complex communication participants were more likely to categorise viable flakes as non-viable. This may be because participants were using a selection method by which they selected the better subset (e.g., half) of the flakes they produced, as if filling a quota of flakes. Although this was disincentivised by the payment method, it is supported by instances in the verbal teaching condition where participants said they did not want to select a flake, not because it was non-viable, but because they already had a number of good flakes and so felt they did not need any more. It is also potentially problematic that participants were not incentivised to correctly identify flakes as participants were rewarded for all viable flakes, regardless of their categorisation. I used this approach as it was quick to carry out, however, given the above, it is clear that a better design would have included penalties for both false negatives and false positives, this approach was used as it was quick to carry out. Although there are some potential insights into the abilities of participants to identify flakes, given the problems described, it is difficult to draw any firm conclusions.

#### **5.6.6 - Individual flakes**

Despite the numerous effects of condition described above there is very little evidence for an effect of condition on any of the parameters of individual flakes. Where differences between conditions are observed they do not show a consistent trend across multiple conditions and so, given the number of analyses carried out, may well be the products of chance. Indeed, control over the size and shape of individual flakes is very difficult skill to master (Callahan, 1979), with other data showing improvement in this regard over a period of decades (Nonaka et al., 2010) - much

longer than the time allowed in this experiment. Given this it seems that the ability of participants to produce individual flakes of a better quality was most likely constrained either because the period of learning allowed was insufficient to transfer the relevant knowledge, or that it was insufficient to allow participants to put relevant knowledge into practice. If the latter were true a longer period of practice - even in the absence of a tutor - would be sufficient to see differences develop, whilst if the former were true an extended period in the presence of the tutor would lead to differences. Thus, this is a question that could be answered with a similar but extended experimental design.

It is also possible, however, that participants were capable of producing individual flakes of a high quality, but were not motivated to do so. The payment system was based on the number of viable flakes, not of their individual quality. However, given the skill required to manipulate the properties of individual flakes I find this interpretation less plausible.

#### **5.6.7 - Effects of position along chains**

Across the numerous analyses there is a general trend for a decrease in performance along chains, particularly in conditions that allowed more complex communication. For example, in VT along chains participants produced fewer viable flakes, a lesser proportion of their produce was viable, they produced more non-viable flakes, they hit more rapidly and the probability that each hit produced a (viable) flake decreased. GT

and BT show a similar though less pronounced pattern. With OL there is no evidence of effects of chain position. With RE there is evidence that some aspects of performance increased along chains; participants made more flakes, more viable flakes, a greater total edge and each hit became more likely to produce a viable flake, although they also produced more non-viable flakes. These increases in performance are hard to interpret. The gradients are typically small and given the number of analyses carried out it is possible that they could be artefacts perhaps driven by a small number of particularly poor participants at the start of chains. However, it is also possible that RE chains converged on a form of knapping that could more effectively be transmitted through observation of results alone, facilitating a slight performance increase along chains. It is also possible that social learning beyond our intentions was able to occur. For example, the noise of striking could be heard through the walls and could have communicated knowledge relevant to striking rate or force. Indeed, from observations of the participants there was a general trend for the force used to increase along chains as participants struggled to make flakes. If participants were able to collect such information from the sounds of knapping through walls then the slight performance increase in RE could be driven by increased force used, increased striking rate or even increased tenacity on the part of the participants. However, it should be noted that high force and striking rate are not part of skilled knapping, I just suggest they might have allowed some small improvement along chains with very limited social transmission.

Given that more complex communication led to better performance, it might seem counterintuitive that these conditions also produced the greatest drop in performance along chains. However, I would argue that it is precisely because of their better

performance that performance decreased along chains. Conditions with complex communication allowed better initial transmission of knapping technology such that there was plenty of scope for a subsequent decrease along chains. More basic conditions, however, resulted in such limited transmission that there was little possibility of further decreases. Further work could investigate means by which the decay in performance could be slowed or stopped as the transmission we observed is clearly insufficient to maintain the behaviours in the long term. An obvious start would be to increase the amount of time participants had to learn the technology, which may well be sufficient to stabilise transmission. However, it is also worth consideration that our chains were only a single participant "wide" in that each participants had access to a single tutor and taught a single pupil, this is not like natural populations where individuals can learn from many individuals, and such narrow chains are known to inhibit stable cultural transmission (Enquist et al., 2010). Thus, a valuable future study could use a "microsociety" structure (Caldwell and Millen, 2008a) to give participants access to multiple tutors.

#### **5.6.8 - Effects of core mass**

The mass of a participant's core had a range of intuitive effects on their performance, allowing participants to make more flakes (both viable and non-viable) and participants with larger cores typically left more of the core unknapped at the end of the test phase. The only unusual finding is that participants with larger cores tended to select a smaller number of flakes than participants with smaller cores. However, given the number of analyses carried out and the difficulty with understanding the criteria

by which participants selected flakes, I do not think this particular result is troubling given that the numerous other effects of core mass found are all highly intuitive.

### **5.6.9 - General Discussion**

Having gone through the various results of this experiment. I shall now consider any general findings that can be brought out of the results as a whole.

#### **5.6.9.1 - The importance of teaching**

These results provide little evidence that even the simplest flaking technology can be easily transmitted through observational learning, despite humans being regarded as particularly adept imitators (Flynn and Whiten, 2008) and other, similarly structured, studies finding that reverse engineering and observational learning is sufficient for cumulative cultural evolution even within 5 minute learning periods (Caldwell and Millen, 2009). Some active or deliberate means of knowledge transfer, for instance through gestural or verbal teaching, seems necessary for any substantive socially mediated increment in knapping performance, at least within the time periods considered in this experiment and none of the performance measures showed an improvement through observing others knapping relative to reverse engineering from previously made tools. There is no evidence, for instance, that observational learning could facilitate an improvement in total flake quality, nor did it increase the rate of

flake production or the proportion of products that were viable flakes. Furthermore, in the non-teaching (RE and OL) conditions, the severe initial drop in performance followed by little further decline suggests that virtually all transmitted knowledge was lost after the first step in the transmission chain, and only the conditions with some form of teaching (BT, GT, VT) showed a more gradual decline and so evidence of information propagation beyond the first step in the chain. In some cases the impact of teaching is quite dramatic with even a small amount of basic tuition (BT) almost doubling the chance that a pupil would produce a viable flake with each hit. Teaching through gesture (GT) supported further increments in flake quality and number, whilst verbal teaching (VT) further increased the rate of production of viable flakes, and the number and quality of flakes produced.

I am not alone in hypothesizing that stone knapping is not a task that lends itself to effective transmission through imitation or emulation (e.g., Stout, 2010). No doubt some information can be acquired this way, for instance, concerning core, hammerstone or flake selection (e.g., Shipton et al., 2009), the requirement to strike the core with the hammerstone, some idea of the force required, and the very idea that flakes can be generated this way. However, the rapid striking action required for effective percussion may inhibit the transmission of more subtle information, for instance, concerning the point of percussion, platform edge and angle, and the precise details of how precisely to hold and hit. Effective transmission of these details may require some form of teaching (i.e. slowing down the striking action to reveal fine-grained aspects of motor patterns, pointing to appropriate targets, demonstrating core rotation, manual shaping of the pupil's grasp).

These findings also support the argument that a complex symbolic language is important for the transmission of the ability to make flakes. It seems plausible that this is because symbolic language enables the transmission of key abstract concepts that may be difficult to transmit otherwise. Despite this suggestion, in this experiment it seems that participants did better with verbal teaching than with gestural teaching, even though both conditions permitted symbolic communication. However, this is consistent with symbolic communication being highly important as it is likely that my participants possessed only a limited symbolic gestural repertoire (e.g., head nodding meaning “yes”). Thus, even though they were allowed to, participants were unable to use gestures as symbols to communicate complex abstract concepts. Supporting this, Ohnuma et al. (1997) commented on the difficulty of communicating the concept of the platform angle without using verbal language. A test of this hypothesis would be to repeat the experiment with participants familiar with sign language, in which case I would expect there to be little difference between verbal and gestural teaching. My findings support the importance of symbolic communication as verbal teaching most frequently resulted in strong evidence of an increase in performance over reverse engineering, but I also find a drop in the amount by which participants talk about the platform angle along chains that mirrors a corresponding decrease in performance. Together, these findings raise the possibility that the importance of language in knapping is that it allows the effective transmission of such concepts. Further experiments could investigate the extent to which participants in non-verbal conditions understand the platform angle.

When considering the importance of language, particularly in contrast to gestural communication, it is important to bear in mind that this study used participants all of

whom were extensively familiar with verbal language, but much less so with gestural communication. This could be problematic as the generally better performance I observed with verbal teaching relative to gestural teaching could be due to participants' lack of familiarity with gestural communication. This is a constraint that is very hard to get around and there is no obvious control. One possibility would be use participants who are familiar with both verbal language and sign language. Should enough individuals be found to carry out this control the difference I observed between verbal and gestural teaching might be expected to disappear. However, it is also worth bearing in mind that by using participants unfamiliar with gestural communication the transmission we observed through gestural teaching could indicate the efficacy of largely non-symbolic gestural communication and so is still of interest.

#### **5.6.9.2 - The co-evolution of stone tools and communication**

A second finding of general significance is that stone tool manufacture would have created a selection pressure favouring increasingly complex gestural teaching and language. To illustrate: Oldowan hominins occupied challenging ecological niches (Blumenshine, 1986; Potts, 2013) in which competition for food resources was strong and (as flint flakes blunt rapidly so cannot be re-used) it is likely that there was a fitness benefit associated with the ability to make and use cutting tools quickly and effectively. Indeed, archaeological evidence shows that efficient core reduction was a goal in the earliest Oldowan (Roche et al., 1999; Semaw, 2000). Such abilities are slow to acquire individually (Callahan, 1979), so social transmission that allowed more rapid acquisition would likely generate fitness benefits. Our data show that



teaching and language are more effective means of transmission of lithic technology than imitation or emulation - allowing participants to produce a greater number and a higher quality of viable flakes from a core, and to do so more efficiently. This implies that the reliance on lithic tools would have created a selection pressure for more teaching and language.

The suggestion that stone toolmaking favoured the evolution of teaching and language in hominins is not new (Uomini, 2009b; Bickerton, 2009; Sterelny, 2012; Stout and Chaminade, 2012), but hitherto has lacked concrete evidence. Additional support for this hypothesis comes from a recent mathematical analysis which found that teaching and cumulative culture would coevolve (Fogarty et al., 2011), experimental studies showing how cumulative cultural learning in children is critically dependent on high-fidelity information transmission (Dean et al., 2012), and theoretical studies showing the importance of high-fidelity transmission for cumulative culture in general (Lewis and Laland, 2012).

Researchers from several disciplines have suggested that this co-evolution supports an autocatalytic processes in which existing technology feeds back to generate selection for further technology and enhanced cognition, including social mechanisms of skill acquisition, eventually leading to exponential growth in human technological complexity (Enquist et al., 2008; Stout, 2011; Uomini, 2009b, 2009a). Consistent with this, there is evidence that, on the broadest scale, Lower Palaeolithic technology is cumulative (Stout, 2011).

In light of these findings, I propose that the widespread use of Oldowan flaking led to selection for more complex forms of social learning involving teaching. This is compatible with the suggestion that the range of variability seen in the early Oldowan (Stout et al., 2010; Hovers, 2012) can be explained by a relatively simple

form of social learning, where imitation/emulation was sufficient to transmit the concept of stone knapping and efficient core maintenance, but the method (sequence of core reduction) was not specifiable by this means, such that each group/individual had its own idiosyncratic core reduction method. The transition to the Acheulean technocomplex at 1.76 mya (Beyene et al., 2013) would indicate that a new form of information transmission was available, which enabled social transmission of the sequence of sub-goals needed to make bifacial technology (Callahan, 1979; Gowlett, 2006).

#### **5.6.9.3 - Implications for Oldowan stasis and the appearance of Acheulean technology**

An outstanding mystery concerning the evolution of stone tools is the apparent stasis of Oldowan flaking technology between 2.5 and 1.5mya (Stout et al., 2010). While Oldowan stone tools are now known to be variable in both space and time, this technology seemingly remains relatively constant for hundreds of thousands of years, an observation for which no satisfactory explanation exists (Roche, 2005; Hovers, 2012; Stout et al., 2010; Stout, 2011). These results imply that in the absence of teaching, Oldowan technology would be unlikely to increase in complexity, because of the limited detail that can be effectively transmitted by observational learning. Even if through extensive practice individuals were to devise more complex reduction strategies, these would be unlikely to be transmitted.

A corollary of this argument is that the sudden appearance of Acheulean technology must be associated with the prior evolution of a new form of social transmission. For example, language is important when learning to make flakes as it enables the transmission of abstract concepts that may be difficult to transmit with gestural communication alone, such as the role of the platform angle in choosing where to strike. In the light of these findings, and given the complexity of the Acheulean and Mousterian Industries, I suggest that the long term stable transmission of these post-Oldowan technologies would indicate the evolution of a symbolic form of communication capable of dealing with abstract concepts, such as language (verbal or gestural). This would suggest that language, at least in some rudimentary form, had evolved by 0.3mya, and perhaps as early as 1.6mya.

#### **5.6.10 - Extensions and improvements**

One limitation of this study is its reliance on a single, 5-minute interaction between pupil and tutor, when it is well established that it takes more than a few hours of practice for modern humans to master even simple flake production, and hundreds of hours for more complex stone tools (Callahan, 1979). This explains the decline along chains observed in the verbal teaching condition, as even with teaching, pupils will require significant practice in order to develop the practical understanding or motor skills alongside theoretical knowledge (Roux et al., 2013; Geribàs et al., 2010). However, the short learning period does not undermine my conclusions, as my focus is not on whether or not it is possible to learn to make stone tools via different mechanisms, but rather on the relative effectiveness of alternative means of social

transmission in the acquisition of such skills. The 5-minute learning period was sufficient to reveal differential rates of skill acquisition and is equivalent in duration to that found in numerous other observational learning experiments, which often find complex skills can be readily acquired through imitation (Flynn and Whiten, 2008). While I cannot rule out the possibility that with more time, participants in the observational learning condition would “catch-up” with those in the verbal teaching condition, it is still of note that I found no evidence of any improvement relative to reverse engineering in the case of observational learning, but did so in the case of the teaching conditions. Indeed, the finding that stone knapping does not lend itself to rapid acquisition through imitation or emulation fits with recent thinking in archaeology (Uomini, 2009b, 2009a; Stout et al., 2011; Stout, 2011). The primary benefit of the short period of learning in this experiment is that it allowed for data collection from a very large number of individuals, with replication across a significant number of conditions, thereby leading to robust findings.

#### **5.6.11 - Conclusions**

In sum, the results presented here find that social transmission, particularly teaching and language, allows individuals to produce tools of greater total utility from a single core and suggest that the longstanding puzzle of stasis in Oldowan Industry likely reflects an inability of hominins to effectively propagate innovations to this lithic technology through observational learning. As a consequence, I suggest that stone toolmaking innovations would be unlikely to spread until hominins evolved sufficient capability to teach the requisite skills. If this is correct, it follows that hominins

developed at least a basic capacity for teaching across the Oldowan. I also present evidence that the appearance of stone technology would have created a selection pressure for increasingly complex communication allowing the more efficient social transmission of the ability to make stone tools leading to a co-evolution of stone tool technology and complex communicative abilities. Given the greater complexity of Acheulian and Mousterian Industries, these findings imply that some form of gestural, or even symbolic, communication was present in order to sustain the stable transmission of such technologies. Together these findings suggest that teaching, and even a form of proto-language, may have evolved substantially earlier than the current consensus view.

## Chapter 6

# **General Discussion**

## 6.1 - Introduction

Over the previous chapters I have presented some empirical data concerning human social learning. In this chapter I shall recap these findings before setting them in a broader context and finally considering areas for further research.

## 6.2 - Summary of findings

The focus of **chapters 3** and **4** was on factors that affected the influence that the decisions of other individuals had on an observer. Regularities in such effects are referred to as "transmission biases" (Boyd and Richerson, 1985) and "social learning strategies" (Laland, 2004) in the cultural evolution literature, or as "trust" (Harris, 2007, 2012) in the developmental psychology literature. Whilst much of the work on these learning rules has involved theoretical modelling (e.g., (Boyd and Richerson, 1985; Kendal et al., 2009; Schlag, 1998; Eriksson et al., 2007), increasingly, theoretical predictions are being tested empirically (Mcelreath et al., 2005; Efferson et al., 2008; Mesoudi and O'Brien, 2008; Caldwell and Millen, 2008b, 2008a; Henrich and Henrich, 2010). In **chapter 3**, I added to this body of findings by investigating whether adult social learning is guided by the number of demonstrators, demonstrator consensus, observer confidence, task difficulty, the cost of asocial learning, task familiarity, demonstrator and observer performance, and their difference and whether the pattern of influence of these factors is in line with predictions from theoretical models.. I found that adults were sensitive to all of these factors, other than their own performance (they instead showed a sensitivity to the demonstrator's performance and the difference between their own performance and that of the demonstrator).

However, a central prediction of theoretical work on this topic is that not only should these biases exist, but that they should also be adaptive in that their use should lead individuals towards acquiring valuable information (Boyd and Richerson, 1985; Laland, 2004; Cavalli-Sforza and Feldman, 1981). This prediction has rarely been empirically tested. In **chapter 3** I show that not only are adults sensitive to a range of factors that influence whether they use social information, but also that the way in which they use social information does act to increase the accuracy of their decision making, satisfying the adaptive predictions of evolutionary models.

In **chapter 4** I presented data on the existence of two of these biases (copy-when-uncertain, and conformist transmission) in young children (ages 3 to 7). I was able to show that the behaviour of the older children was very much like that of adults, showing a greater reliance on social information when the quality of asocial information was low and being disproportionately sensitive to intermediate majorities amongst the informants. The behaviour of younger children was, however, different: they showed little sensitivity to the quality of asocial information and also showed no sensitivity to intermediate majorities, only showing signs of social influence when there was a total consensus amongst the 10 informants. This work not only showed how the form of this behaviour changed as young children aged, but was also able to investigate the extent to which the social influence improved decision making. I found that older children's use of social information was adaptive in that it increased their performance, whilst the younger children's use improved performance to a much lesser extent and only when a total consensus was available. Thus, this data shows



both the development of an adaptive behaviour, but also the development of the adaptiveness of a behaviour.

In **chapter 5** I presented data investigating the role of social learning in the specific case of the development and evolution of lithic tool use in hominins. Using transmission-chain methods established by other cultural evolutionary experiments (Caldwell and Millen, 2008b; Mesoudi and Whiten, 2008; Flynn and Whiten, 2008) I was able to show that (at least in modern humans) teaching and language greatly assist the rapid social transmission of the ability to manufacture Oldowan flint flakes, which first appear in the archaeological record from 2.52mya (Roche et al., 1999; Semaw, 2000). These findings suggest that the million-year stasis observed in the Oldowan technocomplex (Stout et al., 2010) could be due to limitations on the social transmission available to hominins, which may have hindered the stable social transmission of any further developments to this technology. This would imply that the appearance of Acheulean tools (Beyene et al., 2013) indicates the development of more complex means of social transmission, which, I have argued, given the complexity of Acheulean tools, would at least involve some form of teaching, and plausibly a form of verbal instruction. Independent of these possibilities the data also strongly suggest a co-evolution of tools and communication, with tool use creating a selection pressure for more complex means of social transmission which in turn allows for the stable transmission of more complex tool making skills. This data provides empirical support for what previously has remained an often made hypothesis (Uomini, 2009b; Bickerton, 2009; Sterelny, 2012; Stout and Chaminade, 2012), however there are alternative hypotheses. For example, based on data from sub-Saharan Africa and early Upper Palaeolithic Europe it has been argued that

demographic factors such as regional population density and migratory activity impacted upon the development of modern behaviour (Powell et al., 2009). Such a hypothesis has support from work showing that population reduction in Tasmania across 8000 years lead to the loss of several useful technologies (Henrich, 2004b). These hypotheses are not mutually exclusive however, and it is certainly possible that modern human behaviour is the result of both demographic factors and a co-evolution of culture and communicative abilities.

### **6.3 - Broader considerations**

#### **6.3.1 - The status of adaptive learning rules**

The data I presented in chapter 3 provide strong empirical support for the notion of adaptive learning rules that guide reliance on social information such that the fitness of the learner is increased. Evolved learning rules are a central feature of cultural evolutionary theory, distinguishing it from other fields such as social psychology, and so this work provides support for the cultural evolutionary approach (Boyd and Richerson, 1985; Cavalli-Sforza and Feldman, 1981; Laland, 2004). The results presented in chapter 3, along with other empirical work described in Chapters 1 and 2, now provide a detailed account of when adults humans will use social information, and when they will not.

The predictions of cultural evolutionary theory were further tested in the experiment described in Chapter 4 where social learning strategies were investigated in young children. This work highlights how the study of the development of social learning

biases is still possible within an evolutionary framework and clearly shows the relevance of developmental studies of "trust" to cultural evolution theory. Given the above, whilst our knowledge of adaptive learning rules remains far from complete, the general prediction from cultural evolutionary theory, that adaptive learning rules exist and are a general feature of humanity would seem to be well supported.

### **6.3.2 - Conformist Transmission**

Conformist transmission, described in Chapter 2, has been a central focus of the work in this thesis. Whilst I found evidence consistent with it both in adults and in children above the age of 5, elsewhere in the literature it has found more mixed support (Efferson et al., 2008; McElreath et al., 2005). However, it may be possible to reconcile these conflicting findings. For example, in the study by McElreath et al. (2005), which found limited support for conformist transmission, participants were given both asocial and social information, something that theory on conformist transmission has rarely taken into account (Boyd and Richerson, 1985; Henrich and Boyd, 1998). Furthermore, although in Henrich & Boyd's (1998) model individuals are capable of both asocial and social learning, for each decision they choose one source over the other and do not combine the two. In experiments where participants have access to both sources, it seems likely that they would combine the two when making a decision, and not choose one source at the expense of the other. In the work presented here, evidence for conformist transmission, both in adults and children, was found by controlling for asocial information statistically. As McElreath et al. (2005) did not do so this could explain why their data only partially supported conformist transmission.

Furthermore, in Chapter 3, I found that without controlling for asocial information statistically, participant behaviour only matched conformist transmission when the number of demonstrators was high and participant confidence was low. McElreath et al. (2005) used a smaller group of demonstrators than I did (between 3 and 6), which raises the additional possibility that in this study the group sizes studied were too small to observe behaviour consistent with conformist transmission without statistically controlling for asocial information. However, in Efferson et al.'s (2008) study, the social learners did not have access to asocial information, avoiding this problem. In this case, many, but not all, of the participants behaved in line with conformist transmission. The reason why some participants did not conform remains unclear, but plausibly reflects genuine variation in the extent to which individuals are conformist.

Given the above, it is now clear that, at least in some circumstances, human behaviour matches conformist transmission. The next step in this line of research will be to test empirically the predictions of models where conformist transmission was allowed to hinge on other factors, such as the number of traits or the error rate of learning (Nakahashi et al., 2012). Furthermore, the findings of Toelch et al. (2010), make it clear that adults are capable of complex responses to frequency, favouring traits whose popularity is increasing, and researchers should also look to extend theoretical consideration of conformist transmission to incorporate such factors.

### **6.3.3 - Gene-culture co-evolution**

Another distinguishing feature of cultural evolution is the treatment of culture as a partially independent heritable system alongside the more traditional genetic system. The interactions between the two have been the study of many so-called gene-culture co-evolutionary models, and whilst this theory suggests culture can alter the dynamics of genetic evolution, empirical tests of this are difficult. In chapter 5 I provide evidence in support of a co-evolutionary dynamic between the genetic evolution of systems of communication (such as teaching and symbolic communication) and the cultural evolution of stone tools. The proposed reciprocal evolution of technology, part of the cultural inheritance system, and the capacity for social transmission, part of the genetic inheritance system, is accordingly consistent with the dual nature of inheritance emphasised by cultural evolutionary theory. This finding is also of relevance to the ongoing debate as to whether non-genetic inheritance systems (such as culture) are of interest to biologists, with some critics arguing that they tell us nothing biological beyond what can be learned from genes (Dickins and Rahman, 2012). This debate is not limited to the importance of culture, however, and includes other non-genetic inheritance systems such as epigenetics (Jablonka and Lamb, 2006) and niche construction (Odling-Smee et al., 2003). These non-genetic systems have been argued to be involved in various aspects of human evolution, their importance facilitated by our unprecedented capacity for culture and environmental modification (e.g., Shennan, 2011; Gerbault et al., 2011). If evolution is partly a non-genetic process, a stance the data I provide in chapter 5 supports, then these inheritance systems clearly are of interest and importance to biologists. The implications of this debate are considerable, with non-genetic inheritance mechanisms forming part of a

larger movement advocating an extension to the Modern Synthesis (Danchin et al., 2011; Pigliucci and Müller, 2010). Accordingly, studies of gene-culture co-evolution may not only lead to the full incorporation of non-genetic inheritance systems into biology, but may also contribute to the development of a new framework within which to understand the evolutionary process.

## **6.4 - Further work**

### **6.4.1 - The evolved basis of learning rules**

Although the conditions that elicit social learning, and the fact that the human response to these factors broadly adaptive, are reasonably well understood, the evolved mechanistic basis of such learning biases (e.g., genetic, cultural, or otherwise) needs further investigation. Much work on this topic, both theoretical and empirical, has made the phenotypic gambit, assuming that such rules can be effectively treated as under simple genetic control (Laland, 2004; Morgan et al., 2011). Whilst this has been a successful approach, the argument has been made that available evidence suggests such rules are products of cultural evolution, and not genetic evolution (Heyes, 2012). A fuller understanding of social learning requires that this assumption be tested.

Central to this endeavour will be more developmental studies of social learning. These need to identify, not only how social learning strategies develop, but what conditions and other influences affect their development. A good example are studies of social learning in children that differ as to whether they have a secure, ambivalent or

avoidant relationship with their mother at 15 months (Corriveau et al., 2009b). This relationship status, which depends on early interactions between the mother and child, was found experimentally to correlate with the weight children give to their mother's advice at age 5. This shows that the development of adaptive learning rules are at least partially flexible and suggests the phenotypic gambit may be a poor assumption for future work. Moreover, it highlights the possibility that the social learning rules identified so far, largely through work with western undergraduates, may not be universal (Henrich et al., 2010).

Another way in which our understanding of the plasticity in learning rules can be enhanced is through cross cultural studies. Whilst cultural variation is known to exist in reliance on social information (Bond and Smith, 1996) the cause of such variation has not been identified. Experiments with adult participants can identify how much variation exists in social learning and may be able to identify cultural features that correlate with such variation. For example, on the basis of the observation that east-Asian individuals are more reliant on social information than are Western individuals (Bond and Smith, 1996) the prediction could be made that reliance on social information is driven by the holistic thinking style characteristic of east-Asian populations, as opposed to the analytic thinking style of Western populations (Nisbett et al., 2001). Further to this, cross-cultural developmental studies could also identify different developmental trajectories in social learning and could potentially ascertain which experiential factors are involved in the development of transmission biases. Whilst some cross-cultural data on children's use of social information, similar to that in adults, is available (Corriveau and Harris, 2010), it is only able to identify that differences exist, not explain when or how such differences arise.

### **6.4.2 - The neural basis of learning rules**

Whilst the conditions that elicit social learning are well understood, how this behaviour is manifest in the brain is also an area that would benefit from more attention. Chapters 1 and 2 highlight the existing data available on this topic which along with Chapter 3 imply a great variety of strategies are used. Whilst several brain areas have been implicated, much more research is required to understand the role different brain areas play in social decision making. Furthermore, it would be of interest to investigate whether different social learning strategies are the product of different neural mechanisms as opposed to a single decision making mechanism with many different types of input. Likewise, Chapter 4 found differences in social information use as children age, it would be of interest to see what the neural correlates of such development was. Such research could identify if the development of strategies reflected the development of a specific brain area or several. If previous social learning experiments could be replicated with participants within an fMRI machine this would provide a ready source of data that could be understood in the context of the previous work.

### **6.4.3 - Normative motivations for copying**

Almost all of the work on social learning from within cultural evolution has assumed that that main motivation to use social information is to gain access to high quality information (Boyd and Richerson, 1985; Morgan et al., 2011; Mcelreath et al., 2005). Using terminology from social psychology, this is described as an informational motivation (Deutsch and Gerard, 1955). However, experimental cultural evolutionary



studies of the adaptive value of social learning have largely neglected normative influences, where the goal is not high-quality information, but to be in agreement with others (Deutsch and Gerard, 1955). Whilst some theoretical considerations of normative motivations have been made (Boyd and Richerson, 1985) a necessary next step would be to theoretically identify conditions under which such motivations would be expected to evolve, then, as has been done with informationally motivated social learning strategies, these could be tested empirically.

#### **6.4.4 - A quantitative mechanism for social learning**

A long term goal of the study of social learning would be the quantitative description of the mechanism by which individuals combine social and asocial information. Currently we have a firm, qualitative, understanding of this process, such that numerous factors are known to have a broadly positive or negative influence on the likelihood that an individual will adopt the decisions of others. However, a more detailed account of the underlying mechanisms would incorporate not only the general effect, but also its shape (e.g., linear or diminishing returns), and would allow easy comparison across variables (e.g., for given variances, which of two parameters has the greatest influence on decision making). Furthermore, such a mechanism would ideally combine both informational and normative motivations within a single explanatory framework. Previous attempts have been made at such a mechanism from within social psychology (e.g., Social Impact Theory, Latane, 1981; Social Influence Model, Tanford and Penrod, 1984), but none have been entirely successful. However, by combining studies of function from cultural evolution with data from neurobiological studies and by drawing on existing models of asocial decision making

that have proven successful (Gold and Shadlen, 2007) such a goal seems more possible, if still lofty.

## **6.5 - Conclusion**

This thesis set out to experimentally investigate the factors influencing social information use in adults and young children and to examine what forms of social transmission are capable of supporting early stone tool technology. I found that adults on the older children studied showed a great sensitivity to a variety of factors, whilst the youngest children did not. I also found that teaching and language seem very important for the successful transmission of Oldowan technology. I feel this work makes an important contribution to the literature and that the data presented in the previous chapters go some way to answering outstanding questions. However, I also feel the work presented endorses the cultural evolutionary approach to the study of social learning and evolution more generally. Given the above, as well as data from other fields such as social psychology, a thorough understanding of the conditions that elicit social learning is now possible. Additionally, we can be confident that individuals do so adaptively. However, a more detailed account of the mechanism underlying such behaviour is yet to be realised. The numerous successes of cultural evolution and the interdisciplinary interest in social learning hint that such a framework may now be possible and it is certainly an obvious, if daunting, next step.

### Reference List

- Akins, C. K., Klein, E. D., and Zentall, T. R. (2002). Imitative learning in Japanese quail (*Coturnix japonica*) using the bidirectional control procedure. *Anim. Learn. Behav.* 30, 275–81.
- Ambrose, S. H. (2001). Paleolithic Technology and Human Evolution. *Science* (80-. ). 291, 1748–1753.
- Amlacher, J., and Dugatkin, L. (2005). Preference for older over younger models during mate-choice copying in young guppies. *Ethol. Ecol. Evol.*
- Apesteguia, J., Huck, S., and Oechssler, J. (2007). Imitation - theory and experimental evidence. *J. Econ. Theory* 136, 217–235.
- Asch, S. (1956). Studies of independence and conformity. A minority of one against a unanimous majority. *Psychol. Monogr.* 70, Whole No. 416.
- Bardsley, N. (2000). Control Without Deception: Individual Behaviour in Free-Riding Experiments Revisited. *Exp. Econ.* 3, 215–240.
- Baron, R. S., Vandello, J. A., and Brunsman, B. (1996). The forgotten variable in conformity research: Impact of task importance on social influence. *J. Pers. Soc. Psychol.* 71, 915–927.
- Bates, L. a., and Byrne, R. W. (2010). Imitation: what animal imitation tells us about animal cognition. *Wiley Interdiscip. Rev. Cogn. Sci.* 1, 685–695.
- Belfer-Cohen, A., and Goren-Inbar, N. (1994). Cognition and communication in the Levantine Lower Palaeolithic. *World Archaeol.*
- Bentley, R. A., Hahn, M. W., and Shennan, S. J. (2004). Random drift and culture change. *Proc. Biol. Sci.* 271, 1443–50.
- Bentley, R. A., Lipo, C. P., Herzog, H. A., and Hahn, M. W. (2007). Regular rates of popular culture change reflect random copying. *Evol. Hum. Behav.* 28, 151–158.
- Van Bergen, Y., Coolen, I., and Laland, K. N. (2004). Nine-spined sticklebacks exploit the most reliable source when public and private information conflict. *Proc. R. Soc. London Ser. B-Biological Sci.* 271, 957–962.
- Berns, G. S., Capra, C. M., Moore, S., and Noussair, C. (2010). Neural mechanisms of the influence of popularity on adolescent ratings of music. *Neuroimage* 49, 2687–96.

- Berns, G. S., Chappelow, J., Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., and Richards, J. (2005). Neurobiological correlates of social conformity and independence during mental rotation. *Biol. Psychiatry* 58, 245–253.
- Beyene, Y., Katoh, S., Woldegabriel, G., Hart, W. K., Uto, K., Sudo, M., Kondo, M., Hyodo, M., Renne, P. R., Suwa, G., et al. (2013). The characteristics and chronology of the earliest Acheulean at Konso, Ethiopia. *Proc. Natl. Acad. Sci.*
- Bickerton, D. (2009). *Adam's Tongue*. Hill and Wang.
- Biondi, L. M., García, G. O., Bó, M. S., and Vassallo, a. I. (2010). Social Learning in the Caracara Chimango, *Milvago chimango* (Aves: Falconiformes): an Age Comparison. *Ethology* 116, no–no.
- Biro, D., Inoue-Nakamura, N., Tonooka, R., Yamakoshi, G., Sousa, C., and Matsuzawa, T. (2003). Cultural innovation and transmission of tool use in wild chimpanzees: evidence from field experiments. *Anim. Cogn.* 6, 213–23.
- Biro, D., Sousa, C., and Matsuzawa, T. (2006). “Ontogeny and Cultural Propagation of Tool Use by Wild Chimpanzees at Bossou, Guinea: Case studies in nut cracking and leaf folding,” in *Cognitive Development in Chimpanzees*, eds. T. Matsuzawa, M. Tomonaga, and M. Tanaka (Tokyo: Springer).
- Blumenschine, R. J. (1986). *Early hominid scavenging opportunities: implications of carcass availability in the Serengeti and Ngorongoro ecosystems*. Oxford: B. A. R.
- Boesch, C. (1993). “Aspects of transmission of tool-use in wild chimpanzees,” in *Tools, language and cognition in human evolution*, eds. K. Gibson and T. Ingold (Cambridge: Cambridge University Press).
- Boesch, C. (1991). Teaching among wild chimpanzees. *Anim. Behav.* 41, 530–532.
- Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., and White, J.-S. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol. Evol. (Personal Ed.)* 24, 127–35.
- Bonaccio, S., and Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organ. Behav. Hum. Decis. Process.* 101, 127–151.
- Bond, R. (2005). Group size and conformity. *Gr. Process. Intergr. Relations* 8, 331–354.
- Bond, R., and Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) Line judgment task. *Psychol. Bull.* 119, 111–137.
- Botvinick, M. M., Cohen, J. D., and Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn. Sci.* 8, 539–46.

- Boyd, R., and Richerson, P. J. (1988). "An evolutionary model of social learning: the effects of spatial and temporal variation.," in *Social Learning: Psychological and Biological Perspectives*, eds. T. R. Zentall and B. G. Galef (Mahwah, NJ: Lawrence Erlbaum), 29–48.
- Boyd, R., and Richerson, P. J. (1985). *Culture and the Evolutionary Process*. University of Chicago Press
- Braun, D., and Hovers, E. (2009). "Introduction: Current Issues in Oldowan Research," in *Interdisciplinary Approaches to the Oldowan*
- Braun, D. R., and Harris, J. W. K. (2003). Technological developments in the Oldowan of Koobi Fora: Innovative techniques of artifact analysis. *Oldowan rather more than smashing stones*, 117–144.
- Braun, D. R., Plummer, T., Ferraro, J. V., Ditchfield, P., and Bishop, L. C. (2009). Raw material quality and Oldowan hominin toolstone preferences: evidence from Kanjera South, Kenya. *J. Archaeol. Sci.* 36, 1605–1614.
- Braun, D. R., Tactikos, J. C., Ferraro, J. V., Arnow, S. L., and Harris, J. W. K. (2008). Oldowan reduction sequences: methodological considerations. *J. Archaeol. Sci.* 35, 2153–2163.
- Brody, N. (1965). The Effect of Commitment to Correct and Incorrect Decisions on Confidence in a Sequential Decision-Task. *Am J Psychol* 78, 251–256.
- Byrne, R. W. (2009). Animal imitation. *Curr. Biol.* 19, R111–4.
- Byrne, R. W. (2005). "The maker not the tool: the cognitive significance of great ape manual skill," in *Stone Knapping: the necessary conditions for a uniquely hominin behaviour*, eds. V. Roux and B. Bril (Cambridge: McDonald Institute).
- Byrne, R. W., Hobaiter, C., and Klailova, M. (2011). Local traditions in gorilla manual skill: evidence for observational learning of behavioral organization. *Anim. Cogn.* 14, 683–93.
- Caldwell, C. A., and Millen, A. E. (2008a). Experimental models for testing hypotheses about cumulative cultural evolution. *Evol. Hum. Behav.* 29, 165–171.
- Caldwell, C. A., and Millen, A. E. (2009). Social Learning Mechanisms and Cumulative Cultural Evolution: Is Imitation Necessary? *Psychol. Sci.* 20, 1478–1483.
- Caldwell, C. A., and Millen, A. E. (2008b). Studying cumulative cultural evolution in the laboratory. *Philos. Trans. R. Soc. B-Biological Sci.* 363, 3529–3539.
- Callahan, E. (1979). *The basics of biface knapping in the eastern fluted point tradition: A manual for flintknappers and lithic analysts*. Eastern States Archaeological Federation.

- Campbell, J. D., and Fairey, P. J. (1989). Informational and normative routes to conformity: The effect of faction size as a function of norm extremity and attention to the stimulus. *J. Pers. Soc. Psychol.* 57, 457–468.
- Campbell-Meiklejohn, D., Bach, D., and Roepstorff, A. (2010). How the opinion of others affects our valuation of objects. *Curr. Biol.*, 1–6.
- Cavalli-Sforza, L. L., and Feldman, M. W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press
- Chen, E. E., Corriveau, K. H., and Harris, P. L. (2013). Children trust a consensus composed of outgroup members--but do not retain that trust. *Child Dev.* 84, 269–82.
- Clement, F., Koenig, M., and Harris, P. (2004). The Ontogenesis of Trust. *Mind Lang.* 19, 360–379.
- Conolly, J., Colledge, S., Dobney, K., Vigne, J.-D., Peters, J., Stopp, B., Manning, K., and Shennan, S. (2011). Meta-analysis of zooarchaeological data from SW Asia and SE Europe provides insight into the origins and spread of animal husbandry. *J. Archaeol. Sci.* 38, 538–545.
- Coolen, I., van Bergen, Y., Day, R. L., and Laland, K. N. (2003). Species difference in adaptive use of public information in sticklebacks. *Proc. R. Soc. London Ser. B-Biological Sci.* 270, 2413–2419.
- Coolen, I., Ward, A. J. W., Hart, P. J. B., and Laland, K. N. (2005). Foraging nine-spined sticklebacks prefer to rely on public information over simpler social cues. *Behav. Ecol.* 16, 865–870.
- Coolidge, F. L., and Wynn, T. (2005). Working Memory, its Executive Functions, and the Emergence of Modern Thinking. *Cambridge Archaeol. J.* 15, 5–26.
- Corriveau, K. H., Fusaro, M., and Harris, P. L. (2009a). Going with the flow: preschoolers prefer nondissenters as informants. *Psychol. Sci.* 20, 372–7.
- Corriveau, K. H., and Harris, P. L. (2010). Preschoolers (sometimes) defer to the majority in making simple perceptual judgments. *Dev. Psychol.* 46, 437–45.
- Corriveau, K., and Harris, P. L. (2009a). Choosing your informant: weighing familiarity and recent accuracy. *Dev. Sci.* 12, 426–37.
- Corriveau, K., and Harris, P. L. (2009b). Preschoolers continue to trust a more accurate informant 1 week after exposure to accuracy information. *Dev. Sci.* 12, 188–93.
- Corriveau, K., Harris, P., Meins, E., Fernyhough, C., Arnott, B., Elliott, L., Liddle, B., Hearn, A., Vittorini, L., and De Rosnay, M. (2009b). Young children's trust in their mother's claims: Longitudinal links with attachment security in infancy. *Child Dev.* 80, 750–761.

- Coward, F., Shennan, S., Colledge, S., Conolly, J., and Collard, M. (2008). The spread of Neolithic plant economies from the Near East to northwest Europe: a phylogenetic analysis. *J. Archaeol. Sci.* 35, 42–56.
- Danchin, É., Charmantier, A., Champagne, F. a, Mesoudi, A., Pujol, B., and Blanchet, S. (2011). Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. *Nat. Rev. Genet.* 12, 475–86.
- Day, R. (2001). Interactions between shoal size and conformity in guppy social foraging. *Anim. Behav.* 62, 917–925.
- Dean, L. G., Kendal, R. L., Schapiro, S. J., Thierry, B., and Laland, K. N. (2012). Identification of the social and cognitive processes underlying human cumulative culture. *Science* 335, 1114–8.
- Delagnes, A., and Roche, H. (2005). Late Pliocene hominid knapping skills: the case of Lokalalei 2C, West Turkana, Kenya. *J. Hum. Evol.* 48, 435–72.
- Deutsch, M., and Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *J. Abnorm. Soc. Psychol.* 51, 629–636.
- Dickins, T. E., and Rahman, Q. (2012). The extended evolutionary synthesis and the role of soft inheritance in evolution. *Proc. R. Soc. B Biol. Sci.*
- Diez-Martín, F., Sánchez, P., Domínguez-Rodrigo, M., Mabulla, A., and Barba, R. (2009). Were Olduvai Hominins making butchering tools or battering tools? Analysis of a recently excavated lithic assemblage from BK (Bed II, Olduvai Gorge, Tanzania). *J. Anthropol. Archaeol.* 28, 274–289.
- Dindo, M., Whiten, A., and de Waal, F. B. M. (2009). In-Group Conformity Sustains Different Foraging Traditions in Capuchin Monkeys (*Cebus apella*). *PLoS One* 4,
- Drea, C. M., and Wallen, K. (1999). Low-status monkeys “play dumb” when learning in mixed social groups. *Proc. Natl. Acad. Sci. U. S. A.* 96, 12965–9.
- Duffy, G. A., Pike, T. W., and Laland, K. N. (2009). Size-dependent directed social learning in nine-spined sticklebacks. *Anim. Behav.* 78, 371–375.
- Edelson, M., Sharot, T., Dolan, R. J., and Dudai, Y. (2011). Following the crowd: brain substrates of long-term memory conformity. *Science* 333, 108–11.
- Efferson, C., Lalive, R., Richerson, P. J., McElreath, R., and Lubell, M. (2008). Conformists and mavericks: the empirics of frequency-dependent cultural transmission. *Evol. Hum. Behav.* 29, 56–64.
- Engh, A., Esch, K., Smale, L., and Holekamp, K. (2000). Mechanisms of maternal rank “inheritance” in the spotted hyaena, *Crocuta crocuta*. *Anim. Behav.* 60, 323–332.

- Enquist, M., Eriksson, K., and Ghirlanda, S. (2007). Critical social learning: A solution to Rogers's paradox of nonadaptive culture. *Am. Anthropol.* 109, 727–734.
- Enquist, M., Ghirlanda, S., Jarrick, a, and Wachtmeister, C. (2008). Why does human culture increase exponentially? *Theor. Popul. Biol.* 74, 46–55.
- Enquist, M., Strimling, P., Eriksson, K., Laland, K., and Sjostrand, J. (2010). One cultural parent makes no culture. *Anim. Behav.* 79, 1353–1362.
- Eriksson, K., Enquist, M., and Ghirlanda, S. (2007). Critical points in current theory of conformist social learning. *J. Evol. Psychol.* 5, 67–87.
- Eriksson, K., and Strimling, P. (2009). Biases for acquiring information individually rather than socially. *J. Evol. Psychol.* 7, 309–329.
- Flynn, E., and Whiten, A. (2008). Cultural transmission of tool use in young children: A diffusion chain study. *Soc. Dev.* 17, 699–718.
- Fogarty, L., Strimling, P., and Laland, K. N. (2011). The Evolution of teaching. *Evolution (N. Y.)*. 65.
- Forkman, B. (1991). Social facilitation is shown by gerbils when presented with novel but not with familiar food. *Anim. Behav.* 42, 860–861.
- Fusaro, M., and Harris, P. L. (2008). Children assess informant reliability using bystanders' non-verbal cues. *Dev. Sci.* 11, 771–7.
- Fusaro, M., and Harris, P. L. (2013). Dax gets the nod: toddlers detect and use social cues to evaluate testimony. *Dev. Psychol.* 49, 514–22.
- Galef, B., and Whiskin, E. (2008). “Conformity” in Norway rats? *Anim. Behav.* 75, 2035–2039.
- Galef Jr., B. G., Dudley, K. E., and Whiskin, E. E. (2008). Social learning of food preferences in “dissatisfied” and “uncertain” Norway rats. *Anim. Behav.* 75, 631–637.
- Genty, E., Breuer, T., Hobaiter, C., and Byrne, R. W. (2009). Gestural communication of the gorilla (*Gorilla gorilla*): repertoire, intentionality and possible origins. *Anim. Cogn.* 12, 527–46.
- Gerbault, P., Liebert, A., Itan, Y., Powell, A., Currat, M., Burger, J., Swallow, D. M., and Thomas, M. G. (2011). Evolution of lactase persistence: an example of human niche construction. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 366, 863–77.
- Gergely, G., Egyed, K., and Király, I. (2007). On pedagogy. *Dev. Sci.* 10, 139–46.
- Geribàs, N., Mosquera, M., and Vergès, J. M. (2010). What novice knappers have to learn to become expert stone toolmakers. *J. Archaeol. Sci.* 37, 2857–2870.



- Gibson, K., and Ingold, T. (1993). *Tools, language and cognition in human evolution*. Cambridge: Cambridge University Press.
- Gold, J. I., and Shadlen, M. N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–74.
- Gowlett, J. (1979). Complexities of cultural evidence in the Lower and Middle Pleistocene. *Nature* 278.
- Gowlett, J. (2006). “The elements of design form in Acheulian bifaces : modes , modalities , rules and language In press in : The elements of design form in Acheulian bifaces : modes , modalities , rules and,” in *Axe Age: Acheulian Tool-making from Quarry to Discard*, eds. N. Goren-Inbar and G. Sharon (London: Equinox).
- Griskevicius, V., Goldstein, N. J., Mortensen, C. R., Cialdini, R. B., and Kenrick, D. T. (2006). Going along versus going alone: When fundamental motives facilitate strategic (non)conformity. *J. Pers. Soc. Psychol.* 91, 281–294.
- Grüter, C., Czaczkes, T. J., and Ratnieks, F. L. W. (2010). Decision making in ant foragers (*Lasius niger*) facing conflicting private and social information. *Behav. Ecol. Sociobiol.* 65, 141–148.
- Hahn, M. W., and Bentley, R. A. (2003). Drift as a mechanism for cultural change: an example from baby names. *Proc. Biol. Sci.* 270 Suppl , S120–3.
- Halberda, J., and Feigenson, L. (2008). Developmental change in the acuity of the “Number Sense”: The Approximate Number System in 3-, 4-, 5-, and 6-year-olds and adults. *Dev. Psychol.* 44, 1457–65.
- Harris, P. (2012). *Trusting what you’re told: How children learn from others*. Harvard University Press
- Harris, P. L. (2007). Trust. *Dev. Sci.* 10, 135–8.
- Harris, P. L., and Lane, J. Early Learning from Testimony. *Topoi An Int. Rev. Philos.*
- Haun, D. B. M., Rekers, Y., and Tomasello, M. (2012). Majority-biased transmission in chimpanzees and human children, but not orangutans. *Curr. Biol.* 22, 727–31.
- Henrich, J. (2004a). Cultural group selection, coevolutionary processes and large-scale cooperation. *J. Econ. Behav. Organ.* 53, 3–35.
- Henrich, J. (2004b). Demography and Cultural Evolution: How Adaptive Cultural Processes can Produce Maladaptive Losses: The Tasmanian Case. *Am. Antiq.*, 197–214.
- Henrich, J., and Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evol. Hum. Behav.* 19, 215–241.

- Henrich, J., and Gil-White, F. J. (2001). The evolution of prestige - Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evol. Hum. Behav.* 22, 165–196.
- Henrich, J., Heine, S. J., and Norenzayan, A. (2010). Most people are not weird. *Behav. Brain Sci.* 33, 61–83; discussion 83–135.
- Henrich, J., and Henrich, N. (2010). The evolution of cultural adaptations: Fijian food taboos protect against dangerous marine toxins. *Proc. R. Soc. B-Biological Sci.* 277, 3715–3724.
- Henrich, J., and McElreath, R. (2003). The evolution of cultural evolution. *Evol. Anthropol.* 12, 123–135.
- Heyes, C. (2012). Grist and mills: on the cultural origins of cultural learning. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 367, 2181–91.
- Hobaiter, C., and Byrne, R. (2011). The gestural repertoire of the wild chimpanzee. *Anim. Cogn.*, 745–767.
- Hoppitt, W., and Laland, K. (2008). “Social Processes Influencing Learning in Animals: A Review of the Evidence,” in *Advances in the Study of Behavior*, eds. H. Brockmann, T. Roper, M. Naguib, K. Wynne-Edwards, C. Barnard, and J. Mitani, 105–165.
- Horner, V., Proctor, D., Bonnie, K. E., Whiten, A., and de Waal, F. B. M. (2010). Prestige Affects Cultural Learning in Chimpanzees. *PLoS One* 5, -.
- House, B., Henrich, J., Sarnecka, B., and Silk, J. B. (2013). The development of contingent reciprocity in children. *Evol. Hum. Behav.* 34, 86–93.
- Hovers, E. (2012). Invention, reinvention, and innovation: The makings of Oldowan lithic technology. *Orig. Hum. Innov. Creat.* 16, 51–68.
- Jablonka, E., and Lamb, M. (2006). *Evolution in Four Dimensions*. MIT press.
- Jackson, J. M. (1986). In Defense of Social Impact Theory : Comment on Mullen. *J. Pers.* 50, 511–513.
- Jenness, A. (1932). The role of discussion in changing opinion regarding a matter of fact. *J. Abnorm. Soc. Psychol.* 27, 279–296.
- Kameda, T., and Nakanishi, D. (2002). Cost-benefit analysis of social/cultural learning in a nonstationary uncertain environment - An evolutionary simulation and an experiment with human subjects. *Evol. Hum. Behav.* 23, 373–393.
- Kandler, A., and Laland, K. N. (2013). Tradeoffs between the strength of conformity and number of conformists in variable environments. *J. Theor. Biol.* 332C, 191–202.

- Kandler, A., and Shennan, S. (2013). A non-equilibrium neutral model for analysing cultural change. *J. Theor. Biol.* 330, 18–25.
- Kendal, J., Giraldeau, L.-A., and Laland, K. (2009). The evolution of social learning rules: payoff-biased and frequency-dependent biased transmission. *J. Theor. Biol.* 260, 210–9.
- Kendal, R. L., Coolen, I., van Bergen, Y., and Laland, K. N. (2005). Trade-offs in the adaptive use of social and asocial learning. *Adv. Study Behav.* 35, 333–379.
- Key, A. J. M., and Lycett, S. J. (2014). Are bigger flakes always better? An experimental assessment of flake size variation on cutting efficiency and loading. *J. Archaeol. Sci.* 41, 140–146.
- Kim, G., and Kwak, K. (2011). Uncertainty Matters: impact of stimulus ambiguity on infant social referencing. *Infant Child Dev.* 20, 449–539.
- King, A. J., and Cowlshaw, G. (2007). When to use social information: the advantage of large group size in individual decision making. *Biol. Lett.* 3, 137–9.
- Kinzler, K. D., Corriveau, K. H., and Harris, P. L. (2011). Children's selective trust in native-accented speakers. *Dev. Sci.* 14, 106–11.
- Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A., and Fernández, G. (2009). Reinforcement learning signal predicts social conformity. *Neuron* 61, 140–51.
- Koenig, M. a, and Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Dev.* 76, 1261–77.
- Koenig, M. A., Clément, F., and Harris, P. L. (2004). Trust in Testimony: Children's use of true and false statements. *Psychol. Sci.* 15, 694–698.
- Konopasky, R., and Telegdy, G. (1977). CONFORMITY IN THE RAT: A LEADER'S SELECTION OF DOOR COLOR VERSUS A LEARNED DOOR-COLOR DISCRIMINATION. *Percept. Mot. Skills* 44, 31–37.
- Kruschke, J. K. (2010). What to believe: Bayesian methods for data analysis. *Trends Cogn. Sci.* 14, 293–300.
- De la Torre, I. (2011). The origins of stone tool technology in Africa: a historical perspective. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 366, 1028–37.
- Laland, K. N. (2004). Social learning strategies. *Learn. Behav.* 32, 4–14.
- Laland, K. N., Sterelny, K., Odling-Smee, J., Hoppitt, W., and Uller, T. (2011). Cause and Effect in Biology Revisited: Is Mayr's Proximate-Ultimate Dichotomy Still Useful? *Science* (80-. ). 334, 1512–1516.
- Latane, B. (1981). The Psychology of Social Impact. *Am. Psychol.* 36, 343–356.

- Latane, B., and Wolf, S. (1981). The Social Impact of Majorities and Minorities. *Psychol. Rev.* 88, 438–453.
- Lewis, H. M., and Laland, K. N. (2012). Transmission fidelity is the key to the build-up of cumulative culture. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 367, 2171–80.
- Likert, R. (1932). A technique for the measurement of attitudes. *Arch. Psychol.*
- Luna, K., and Martín-Luengo, B. (2012). Confidence-Accuracy Calibration with General Knowledge and Eyewitness Memory Cued Recall Questions. *Appl. Cogn. Psychol.* 26, 289–295.
- Lunn, D., and Spiegelhalter, D. (2009). The BUGS project: Evolution, critique and future directions. *Stat. ... Statistics*, 3049–3067.
- Maestripieri, D. (2007). *Macchiavellian Intelligence: How rhesus macaques and humans have conquered the world*. Chicago/London: University of Chicago Press.
- Mason, M. F., Dyer, R., and Norton, M. I. (2009). Neural mechanisms of social influence. *Organ. Behav. Hum. Decis. Process.* 110, 152–159.
- Mcelreath, R., Lubell, M., Richerson, P., Waring, T., Baum, W., Edsten, E., Efferson, C., and Paciotti, B. (2005). Applying evolutionary models to the laboratory study of social learning. *Evol. Hum. Behav.* 26, 483–508.
- McPherron, S. P., Alemseged, Z., Marean, C. W., Wynn, J. G., Reed, D., Geraads, D., Bobe, R., and Bérarat, H. a (2010). Evidence for stone-tool-assisted consumption of animal tissues before 3.39 million years ago at Dikika, Ethiopia. *Nature* 466, 857–60.
- Mesoudi, A. (2011). An experimental comparison of human social learning strategies: payoff-biased social learning is adaptive but underused.
- Mesoudi, A. (2008). An experimental simulation of the “copy-successful-individuals” cultural learning strategy: adaptive landscapes, producer-scrounger dynamics, and informational access costs. *Evol. Hum. Behav.* 29, 350–363.
- Mesoudi, A., and O’Brien, M. J. (2008). The cultural transmission of Great Basin projectile-point technology I: An experimental simulation. *Am. Antiq.* 73, 3–28.
- Mesoudi, A., and Whiten, A. (2008). Review. The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 363, 3489–501.
- Morgan, T., Rendell, L., Ehn, M., Hoppitt, W., and Laland, K. (2011). The evolutionary basis of human social learning. *Proc. R. Soc. B Biol. Sci.*
- Mullen, B. (1986). Effects of strength and immediacy in group contexts: Reply to Jackson. *J. Pers. Soc. Psychol.* 50, 514–516.

- Mullen, B. (1985). Strength and Immediacy of Sources - a Meta-Analytic Evaluation of the Forgotten Elements of Social Impact Theory. *J. Pers. Soc. Psychol.* 48, 1458–1466.
- Nakahashi, W., Wakano, J. Y., and Henrich, J. (2012). Adaptive social learning strategies in temporally and spatially varying environments: how temporal vs. spatial variation, number of cultural traits, and costs of learning influence the evolution of conformist-biased transmission, payoff-biased transmissi. *Hum. Nat.* 23, 386–418.
- Nicol, C., and Pope, S. (1994). Social learning in small flocks of laying hens. *Anim. Behav.*
- Nisbett, R., Peng, K., Choi, I., and Norenzayan, A. (2001). Culture and systems of thought: holistic versus analytic cognition. *Psychol. Rev.*, 291–310.
- Nonaka, T., Bril, B., and Rein, R. (2010). How do stone knappers predict and control the outcome of flaking? Implications for understanding early stone tool technology. *J. Hum. Evol.* 59, 155–167.
- Nowak, A., Szamrej, J., and Latane, B. (1990). From Private Attitude to Public-Opinion - a Dynamic Theory of Social Impact. *Psychol. Rev.* 97, 362–376.
- Odic, D., Hock, H., and Halberda, J. (2012). Hysteresis Affects Approximate Number Discrimination in Young Children. *J. Exp. Psychol. Gen.*
- Odling-Smee, F. J., Laland, K. N., and Feldman, M. W. (2003). *Niche Construction: The neglected process in evolution*. Princeton University Press.
- Ohnuma, K., Aoki, K., and Akazawa, T. (1997). Transmission of Tool-making through Verbal and Non-verbal Communication--Preliminary Experiments in Levallois Flake Production. *Anthropol. Sci.* 105, 159–168.
- Pasquini, E. S., Corriveau, K. H., Koenig, M., and Hanis, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Dev. Psychol.* 43, 1216–1226.
- Pigliucci, M., and Müller, G. B. (2010). *Evolution: The Extended Synthesis*. MIT press.
- Pike, T., Kendal, J., Rendell, L., and Laland, K. N. (2010). Learning by proportional observation in a species of fish. *Behav. Ecol.*, 1–18.
- Pike, T. W., and Laland, K. N. (2010). Conformist learning in nine-spined sticklebacks' foraging decisions. *Biol. Lett.* 6, 466–8.
- Pleskac, T. J., and Busemeyer, J. R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychol. Rev.* 117, 864–901.
- Potts, R. (2013). Hominin evolution in settings of strong environmental variability. *Quat. Sci. Rev.* 73, 1–13.

- Powell, A., Shennan, S., and Thomas, M. G. (2009). Late Pleistocene demography and the appearance of modern human behavior. *Science* 324, 1298–301.
- Reader, S. (2000). Social Learning and Innovation: Individual differences, diffusion dynamics and evolutionary issues.
- Rendell, L., Fogarty, L., Hoppitt, W. J. E., Morgan, T. J. H., Webster, M. M., and Laland, K. N. (2011). Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends Cogn. Sci.* 15, 68–76.
- Rendell, L., Fogarty, L., and Laland, K. N. (2010). Rogers' paradox recast and resolved: population structure and the evolution of social learning strategies. *Evolution* 64, 534–48.
- Richerson, P. J., and Boyd, R. (2005). *Not By Genes Alone: How Culture Transformed Human Evolution*. University Of Chicago Press.
- Roche, H. (2005). "From simple flaking to shaping: stone-knapping evolution among early hominins," in *Stone Knapping: the necessary conditions for a uniquely hominin behaviour*, eds. V. Roux and B. Bril (Cambridge: McDonald Institute).
- Roche, H., Delagnes, a, Brugal, J. P., Feibel, C., Kibunjia, M., Mourre, V., and Texier, P. J. (1999). Early hominid stone tool production and technical skill 2.34 Myr ago in West Turkana, Kenya. *Nature* 399, 57–60.
- Rogers, A. R. (1988). Does Biology Constrain Culture. *Am. Anthropol.* 90, 819–831.
- Roux, V., Bril, B., and Dietrich, G. (2013). learning Skills and in stone involved of stone-bead case India difficulties knapping : knapping the in. 27, 63–87.
- Schlag, K. H. (1999). Which one should I imitate? *J. Math. Econ.* 31, 493–522.
- Schlag, K. H. (1998). Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. *J. Econ. Theory* 78, 130–156.
- Semaw, S. (2000). The World's Oldest Stone Artefacts from Gona, Ethiopia: Their Implications for Understanding Stone Technology and Patterns of Human Evolution Between 2.6–1.5 Million Years Ago. *J. Archaeol. Sci.* 27, 1197–1214.
- Shennan, S. (2011). Descent with modification and the archaeological record. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 366, 1070–9.
- Shennan, S., and Wilkinson, J. (2001). Ceramic style change and neutral evolution: a case study from Neolithic Europe. *Am. Antiq.* 66, 577–593.
- Shepard, R. N., and Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science* (80-. ). 171, 701.
- Sherif, M. (1935). A study of some social factors in perception. *Arch. Psychol.* 187, 60.

- Shipton, C. B. K., Petraglia, M. D., and Paddayya, K. (2009). Stone tool experiments and reduction methods at the Acheulean site of Isampur Quarry , India. 83, 769–785.
- Smith, C. E., Blake, P. R., and Harris, P. L. (2013). I should but I won't: why young children endorse norms of fair sharing but do not follow them. *PLoS One* 8, e59510.
- Sterelny, K. (2012). Language, gesture, skill: the co-evolutionary foundations of language. ... *Trans. R. Soc. B* ..., 2141–2151.
- Stout, D. (2010). Possible relations between language and technology in human evolution. *Stone tools Evol. Hum.* ....
- Stout, D. (2011). Stone toolmaking and the evolution of human culture and cognition. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 366, 1050–9.
- Stout, D., and Chaminade, T. (2012). Stone tools, language and the brain in human evolution. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 367, 75–87.
- Stout, D., Passingham, R., Frith, C., Apel, J., and Chaminade, T. (2011). Technology, expertise and social cognition in human evolution. *Eur. J. Neurosci.* 33, 1328–38.
- Stout, D., Semaw, S., Rogers, M., and Cauche, D. (2010). Technological variation in the earliest Oldowan from Gona, Afar, Ethiopia. *J. Hum. Evol.* 58, 474–491.
- Stout, D., Toth, N., Schick, K., Stout, J., and Hutchins, G. (2000). Stone tool-making and brain activation: position emission tomography (PET) studies. *J. Archaeol.*
- Tamis-LeMonda, C. S., Adolph, K. E., Lobo, S. a, Karasik, L. B., Ishak, S., and Dimitropoulou, K. a (2008). When infants take mothers' advice: 18-month-olds integrate perceptual and social information to guide motor action. *Dev. Psychol.* 44, 734–46.
- Tanaka, M. M., Kendal, J. R., and Laland, K. N. (2009). From Traditional Medicine to Witchcraft: Why Medical Treatments Are Not Always Efficacious. *PLoS One* 4, -.
- Tanford, S., and Penrod, S. (1984). Social-Influence Model - a Formal Integration of Research on Majority and Minority Influence Processes. *Psychol. Bull.* 95, 189–225.
- Toelch, U., Bruce, M. J., Meeus, M. T. H., and Reader, S. M. (2010). Humans copy rapidly increasing choices in a multiarmed bandit problem. *Evol. Hum. Behav.* 31, 326–333.
- Tong, E. M. W., Tan, C. R. M., Latheef, N. A., Selamat, M. F. B., and Tan, D. K. B. (2008). Conformity: Moods matter. *Eur. J. Soc. Psychol.* 38, 601–611.

- Torre, I. de la (2011). The origins of stone tool technology in Africa: a historical perspective. *Philos. Trans. ...*, 1028–1037.
- Toth, N. (1987). Behavioral inferences from early stone artifact assemblages: an experimental model. *J. Hum. Evol.*
- Uomini, N. T. (2009a). “Prehistoric left-handers and prehistoric language,” in *The emergence of cognitive abilities: the contribution of neuropsychology to archaeology*, eds. S. A. de Beaune and F. L. Coolidge (Cambridge, UK: Cambridge University Press), 37–55.
- Uomini, N. T. (2009b). The prehistory of handedness: archaeological data and comparative ethology. *J. Hum. Evol.* 57, 411–9.
- Visalberghi, E., and Frigaszy, D. (1995). The behaviour of capuchin monkeys, *Cebus apella*, with novel food: the role of social context. *Anim. Behav.* 49, 1089–1095.
- Weizsäcker, G. (2010). Do we follow others when we should? A simple test of rational expectations. *Am. Econ. Rev.* 100, 2340–2360.
- Whiten, A., Horner, V., and de Waal, F. B. M. (2005). Conformity to cultural norms of tool use in chimpanzees. *Nature* 437, 737–740.
- Wood, L. a., Kendal, R. L., and Flynn, E. G. (2012). Context-dependent model-based biases in cultural transmission: children’s imitation is affected by model age over model knowledge state. *Evol. Hum. Behav.* 33, 387–394.
- Wynn, T., Hernandez-Aguilar, R. A., Marchant, L. F., and McGrew, W. C. (2011). “An ape’s view of the Oldowan” revisited. *Evol. Anthropol.* 20, 181–97.
- Yaniv, I. (2004). The Benefit of Additional Opinions. *Curr. Dir. Psychol. Sci.* 13, 75–78.