

Cognitive Capacities Underlying Cumulative Culture: A Developmental Approach

Charlotte Elizabeth Holmes Wilks

Thesis submitted for the degree of Doctor of Philosophy

December 2020

University of Stirling

**UNIVERSITY of
STIRLING**



Abstract

In this thesis I examine cognitive capacities which may explain human propensity for cumulative culture, and its near absence in nonhumans. Although the potential importance of human-unique cognition has been discussed in the literature, the particular capacities responsible remain unknown. Adults show cumulative culture, therefore any capacities implicated are expected to develop during childhood. I thus take a developmental approach (in children aged 3-10): across four experimental studies, I investigate whether children's use of social information, and thus potential for cumulative culture, changes in line with the development of human-unique cognitive capacities. The first research strand utilised a novel experimental approach to examine the cognitive constraints on utilising social information in ecologically valid contexts. The results highlighted the importance of both task context and cognitive ability for cumulative culture. We thus propose that human cognition may enable our species to accumulate culture across a broader range of contexts than nonhumans, including utilising information from multiple social models. Furthermore, based on the contexts presented, we found that general cognitive abilities, such as working memory (neglected in much of the literature, in favour of social propensities), may be important for human cumulative culture. A second strand of the thesis focused on human-unique social cognitive abilities, examining whether these enable cumulative culture through the selective copying of more effective individuals/trait variants. Specifically, we hypothesised that the development of an explicit understanding of others' minds may coincide with a shift in the flexibility of social learning strategies – from nonhuman-like, associative strategies to adult-like, understanding-based strategies. However, we did not identify such a shift in the age range tested (3-8 years). In conclusion, the cognitive capacities underlying human cumulative culture were not definitively determined, yet this thesis has furthered understanding of the general and social cognitive capacities which may be implicated.

Acknowledgements

Firstly, I thank my partner Timothy McGowan, for being there, putting up with my nocturnal thesis-writing habits and always believing that I am capable of achieving a PhD. I am equally thankful to my parents: My mother, Christine Wilks, for encouraging me to pursue my interests and for being the most patient person I have ever met. My father, Peter Wilks, for instilling in me the drive to avoid a “dead-end job” (perhaps non-standard, but effective). I am grateful for the sacrifices you have both made for me.

On a lighter, but nevertheless important note, my pet rabbit Puddles (and her late companion Maisey) have offered unconditional and honest friendship throughout my PhD. This has been a huge source of comfort.

Having aspired to undertake a PhD for a number of years, being part of the Ratchecog team at Stirling did not disappoint. My colleagues, and friends, have inspired me to be the best scientist that I can be and have offered continued support. Thank you: Mark Atkinson, Kirsten Blakey, Donna Kean, Gemma Mackintosh, Elizabeth Renner and Juliet Robertstone (and my supervisor Christine Caldwell – I’ll continue my thanks to you below).

My additional supervisors (Eva Rafetseder, Elizabeth Renner, and honorary supervisor Mark Atkinson) have provided essential knowledge and advice, both in relation to the PhD studies and academia in general. Thank you for taking the time out of your busy schedules to share your expertise; I hope that I can play as supportive a role for an aspiring academic in the future!

I would like to especially thank my supervisor, Christine Caldwell. Working as a teacher, I had *almost* given up on my PhD dream when you gave me this opportunity. Whatever I go on to do, in this uncertain environment, I feel very lucky to have been given these four years to explore such an interesting topic and to think about concepts I still don’t feel qualified to think about. I am extremely fortunate to have had a supervisor who offered both guidance and freedom and, above all else, one who continues to inspire me.

Finally, the research I conducted would not have happened without the participation of some 600 plus children and the wide-range of organisations which made this possible. I therefore thank these children and their parents: Children – I hope you enjoyed my puppeteering skills, and parents – I hope you benefitted from the brief respite offered. Thank you to Glasgow Science Centre, Blair Drummond Safari Park and The Scout Association for supporting my research, and to the members of staff and volunteers who have gone out of their way to accommodate the studies. Thank you also to the nurseries, schools and teachers, in Stirling and Bradford, who showed enthusiasm and support for this project – I

know that your time is limited! The University of Stirling, Psychology Kindergarten was an especially essential resource which I have turned to throughout my PhD. I appreciate the help and enthusiasm I have received from the Kindergarten staff, in particular Federica Caruso and Sarah Thorburn.

Contents

Chapter 1: General Introduction	9
Culture, Cumulative Culture and the Ratchet Effect	9
Proposals for the Uniqueness of Cumulative Culture in Humans.....	11
Social Cognitive Abilities and Social Learning Mechanisms	12
Social Learning Strategies and Explicitly Metacognitive Social Learning Strategies	14
Physical Intelligence and the Technical Reasoning Hypotheses.....	17
Demography.....	18
Brief Review of Experimental Methods Used to Investigate Cumulative Culture.....	19
Generational Replacement Methods	19
Closed Group	21
An Alternative Method	22
Studies of Cumulative Culture in Children	26
Contribution of This Thesis.....	30
Chapter 2: Cognitive Prerequisites for Cumulative Culture are Context-Dependent: Children’s Potential for Ratcheting Depends on Cue Longevity.....	34
Abstract.....	34
Introduction	35
Materials and Methods.....	44
Participants	44
Apparatus.....	44
Procedure	45
Puppet Selections.....	46
Memory Questions.	46
Child Selections.....	47
Results.....	47
Optimal Response Count	48
Repeating.....	50
Effects of Information Type, Age, and Condition.....	51
Effects of Information Type and Age in the Enduring Cues and Transient Information Conditions.	52
Effects of Age and Condition by Information Type.....	54
Potential for Ratcheting (PFR).....	55
Discussion.....	60
Chapter 3: Children’s Use of Social Information From Multiple Models: Cognitive Capacities Underlying Population Size Effects on Cumulative Culture	64
Abstract.....	64
Introduction	65

Measuring the Potential for Cumulative Culture.....	75
Materials and Methods.....	76
Participants	76
Experiment.....	77
Demonstration Types.	78
Procedure	79
Introduction to Task.....	79
Practice Trials (x1 Per Condition).	81
Experimental Trials (x 6 Per Condition).	84
Ending of Experiment.	84
Results.....	84
Score.....	85
Score – Effects of Age and Condition.	85
Score – Effects of Age by Condition.	86
Score – Effects of Condition by Age.	87
Repeating.....	87
Rewarded Repeats.....	87
Unrewarded Repeats.....	92
Potential for Cumulative Culture	92
Discussion.....	97
Chapter 4: Who Knows? Does Children's Developing Understanding of Others' Minds Affect Their Social Learning Strategies? Part 1.....	103
Introduction	103
Method.....	110
Participants	110
Materials	112
Procedure	113
Social Information Trial.....	114
Test Trial.	115
Results.....	116
Basket Choices	116
Verbal Responses.....	124
Discussion.....	130
Chapter 5: Who Knows? Does Children's Developing Understanding of Others' Minds Affect Their Social Learning Strategies? Part 2.....	135
Introduction	135
Method.....	139
Participants	139
Materials	141
Procedure	141

Social Information Trials.....	142
Test Trials.....	143
Results.....	144
Basket Choices.....	144
Verbal Responses.....	156
Discussion.....	159
Chapter 6: General Discussion.....	165
Summary of Thesis Aims and Hypotheses.....	165
Summary and Discussion of the Work Presented in Chapters 2-5.....	166
Chapter 2.....	166
Chapter 3.....	168
Chapters 4 and 5.....	171
Limitations and Future Directions.....	174
Final Comments.....	176
References.....	179
Appendix.....	195
Supplementary Information – Chapter 2.....	195
Materials and Methods.....	195
Participants – Glasgow Science Centre.....	195
Participants – School.....	195
Results.....	195
Optimal Response Count.....	195
Repeating.....	195
Potential for Ratcheting.....	196
Memory Question.....	197
Figures.....	201
Supplementary Information – Chapter 3.....	202
Results.....	202
Rewarded Repeats – Effects of Age by Condition.....	202
Rewarded Repeats – Effects of Condition by Age.....	203
Figures.....	208
Experimenter Verbal Scripts – Chapters 2-5.....	208

Tables and Figures

	Chapter 1	Chapter 2	Chapter 3	Chapter 4	Chapter 5	Appendix
Figures	1.1	2.1	3.1	4.1	5.1	S.1
		2.2	3.2	4.2	5.2	S.2
		2.3	3.3	4.3	5.3	S.3
		2.4	3.4			
		2.5	3.5			
			3.6			
			3.7			
			3.8			
Tables		2.1	3.1	4.1	5.1	S.1
		2.2	3.2	4.2	5.2	S.2
				4.3	5.3	S.3
				4.4	5.4	S.4
				4.5	5.5	S.5
					5.6	
					5.7	

Chapter 1: General Introduction

Culture, Cumulative Culture and the Ratchet Effect

As humans we inhabit a vast cultural world. Simply sitting here at my desk, I am surrounded by objects and technologies (a computer, a printer, the cars through the window) which have resulted from a process of continued improvement by the hands and minds of many. As a researcher I use language to discuss my ideas and I operate within societal systems which have evolved over generations. Such products and processes, also termed cultural traits, develop to become more suited for purpose, more functional. They are ubiquitous in human populations and are considered examples of cumulative culture (or, equally, cumulative cultural evolution). I now provide a more formal definition of these terms, one which is assumed throughout this thesis and would be considered accurate by most scientists in our field (see Mesoudi & Thornton, 2018, for further discussion regarding how cumulative culture is defined).

Cumulative culture is a process whereby a directional pattern of change results in “improvements” (Tennie et al., 2009) or increasingly “preferred” traits (Caldwell, 2018) over generations of social learners (Boyd & Richerson, 1996; Mesoudi & Thornton, 2018; Tomasello, 1999). This definition is in line with four core criteria, recently outlined by Mesoudi and Thornton (2018) and based on a sampling of over 30 publications, which are as follows: i) a change to a behaviour or cultural product, ii) social transmission of the modified trait, iii) improvement in performance as a consequence of the modification, and iv) iteration of these steps resulting in ongoing improvement over time. These criteria were common across the entire sample and thus can be considered a consensus definition. However, some researchers adhere to definitions which are more restrictive, or encompass additional requirements (e.g. Enquist et al., 2011; Hunt & Gray, 2003; Pradhan et al., 2012; Reindl et al., 2020), as captured in Mesoudi and Thornton's (2018) “extended criteria”. Furthermore, the notion of constant improvement has led to cumulative culture also being referred to as the *ratchet effect* (Tennie et al., 2009; Tomasello et al., 1993). A ratchet is a mechanical device which allows motion in one direction only, thus similarly the ratchet effect (or “ratcheting”) describes continuing improvement to a cultural trait without backwards slippage to a less functional form. I use the terms ratchet effect, ratcheting and cumulative culture interchangeably.

Human propensity for cumulative culture has enabled us to dominate the planet, benefiting from the achievements of others in a manner seemingly unparalleled in nonhuman animals (henceforth nonhumans). Therefore, understanding cumulative culture may be fundamental to understanding what it means to be human. Yet, elucidating what may be

unique about human culture – and why it is such a wide-spread phenomenon – necessarily involves consideration of the impressive cultural abilities of nonhumans. There are increasingly more published examples of *culture* in nonhumans (Laland & Hoppitt, 2003; Whiten et al 1999), yet very few examples of *cumulative* culture (Dean et al., 2014). Setting the somewhat controversial subject of nonhuman cumulative culture aside for a moment, I define a culture according to Laland and Hoppitt (2003): a behaviour transmitted between individuals within a population (via social learning – see Reader & Biro, 2010) which becomes readily used by most members of the population (i.e. is a “group-typical behaviour pattern”) yet is absent in other populations of the same species. Unlike an example of cumulative culture, a culture does not necessitate a change or improvement to an existing behaviour or cultural product – it does not necessarily involve ratcheting. There is some disagreement amongst scientists as to which nonhuman species can be considered to display culture (Dean et al., 2014; Laland & Galef, 2009), with this depending on the definition used and the quality of evidence demanded. For example, in the past there has been debate over whether culture is restricted to primates (Whiten et al., 1999; Whiten & van Schaik, 2007), or present in other mammals, birds and fish (Laland & Hoppitt, 2003). Nonetheless, it is now clear that culture is far from an exclusively human phenomenon (Dean et al., 2014; Whiten et al., 2016) and examples, which adhere to the above definition, have now been confirmed in a range of diverse species, including great tits (Aplin et al., 2015), humpback whales (Allen et al., 2013; Garland et al., 2011) and chimpanzees (Hobaiter et al., 2014; Luncz & Boesch, 2014).

On the other hand, *cumulative* culture has been proposed to be unique to humans (Tomasello et al., 1993) and its existence in nonhumans is contentious (Dean et al., 2014). Additional evidence is needed before we can definitively confirm or deny whether nonhumans are capable of cumulative culture, and the answer is likely to be more nuanced (and, as in the case of culture, dependent on the definition used). However, recent research (outlined briefly below) raises the possibility that some animal behaviours may fulfil the requirements for cumulative culture under restricted circumstances (a full review of the nonhuman literature is beyond the scope of this thesis).

In line with the above criteria for cumulative culture (Mesoudi & Thornton, 2018) an experimental study by Sasaki and Biro (2017) showed that the flight routes of homing pigeons became more efficient over generations of birds flown together (although some scientists would question whether increasing efficiency should constitute cumulative culture, see Wilks & Blakey, 2018, for a discussion on this). These authors provided evidence that a process of collective intelligence, in which individuals within a population pool information, can generate cumulative improvements to shared behaviours which are repeated over

multiple attempts.

Two field studies (Jesmer et al., 2018; Schofield et al., 2017) now claim to evidence cumulative culture in real-world, nonhuman populations. Jesmer et al. (2018) studied the adaptive migration routes followed by bighorn sheep and moose that had been translocated into unfamiliar environments. Compared to historic populations, fewer translocated sheep adaptively migrated towards high quality forage. Furthermore, there was a positive association between this behaviour and length of time since introduction into an environment in populations of sheep and moose that had been translocated 10 to 110 years previously. The authors thus proposed that adaptive migration routes are culturally transmitted and represent an example of ratcheting – sheep learn migration routes via social learning and these are honed over generations of learners. However, this proposal requires the inference that historical populations exhibited ratcheting based on differences between these and translocated populations, thus providing only circumstantial evidence for cumulative culture (Caldwell et al., 2020). The other, aforementioned field study (Schofield et al., 2017) examined 60-years-worth of ethnographic data from Japanese macaques in Koshima who had been provisioned with unwashed sweet potatoes and unhusked wheat grains. The monkeys exhibited washing behaviours which the authors claim to have undergone a process of cumulative improvement. However, there was no evidence that the behaviours were socially transmitted between individuals – a key requirement for cumulative culture (Mesoudi & Thornton, 2018). Instead, the changes to these behaviours may have been attributable to other factors, such as patterns of provisioning or the availability of alternative food sources.

The above nonhuman studies open up the possibility that instances of ratcheting may be found across more taxa than has been previously assumed and that cumulative culture could, in some instances, occur in the absence of complex cognitive abilities (Sasaki & Biro, 2017). However, based on research to date it seems that cumulative culture in nonhumans is likely restricted in the range of contexts for which it occurs in any one species, and examples are likely to be of a simpler form to that we see in human societies. We consider effects of task context on ratcheting in Chapter 2.

Proposals for the Uniqueness of Cumulative Culture in Humans

Even if we acknowledge that nonhumans may show examples of cumulative culture under restricted circumstances, it is undeniable that there is something unique about the phenomenon as observed in humans. This distinctiveness of human cumulative culture has captured the attention of scientists across a variety of disciplines and I will now summarise the most prominent explanations. I then move on to discuss the experimental methods which

have been utilised in the investigation of cumulative culture (pp. 19-26) and to outline experimental work focussed on human children (henceforth children), pp. 26-30. This thesis aims to add to scientific knowledge in this regard (pp. 30-34) .

Social Cognitive Abilities and Social Learning Mechanisms

Multiple theories have highlighted that humans may use social information differently to nonhumans due to unique or enhanced social cognitive abilities and/or social learning mechanisms (Boyd & Richerson, 1996; Dean et al., 2014; Tomasello, 1990, 1999; Tomasello et al., 1993, 1994). Tomasello et al. (1993) has drawn a distinction between “social learning”, which is said to occur in both humans and many nonhuman species, and “cultural learning”, proposed to be dependent on human-unique aspects of social cognition and to have evolved as an adaptation to social problems in human societies (e.g. a need for communication, cooperation, competition or deception). In social learning the learning process itself is not inherently social, but the social environment allows the developing individual to become exposed to information which they can then learn individually e.g. a young chimpanzee having their attention drawn towards the nut, hard surface and rock-hammer required for successful nut cracking through the nut-cracking actions of its mother. However, according to this account, human-unique cultural learning is different in that an individual learns a new behaviour through their observation of another individual performing the behaviour.

Therefore, social cognitive abilities which allow understanding of this individual’s mind are proposed to be beneficial and to underlie social learning mechanisms possessed only by humans. Such human-unique mechanisms include imitative, instructed and collaborative learning, the use of which are considered to result in the cumulative aspect of human culture through allowing information to pass between individuals with high fidelity. Correspondingly, it has been argued by Tennie et al. (2009) that nonhumans’ more limited social learning mechanisms (particularly, a focus on product copying rather than imitation and teaching) prevent cumulative culture by limiting nonhumans to their species’ physical cognition skills i.e. their “zone of latent solutions”. Tennie et al. (2009) posit that this is the reason nonhuman cultural behaviours do not become more complex than those which could theoretically be invented by one individual of the species, under suitable environmental conditions.

Related to the idea of cultural learning is the “cultural intelligence hypothesis” (Herrmann et al., 2007) which is also centered around the notion that humans have enhanced social cognitive skills in comparison to nonhumans, specifically in social learning, communication and theory of mind. These skills are said to have evolved in response to

“ultra-social” living, enabling humans to create cultural groups within which distinctive cultural practices, artefacts and tools are established; children’s specialised socio-cognitive skills are said to develop early and allow absorption into their cultural world through participation in cultural practices and the use of tools/artefacts. More recently, Dean et al. (2012, 2014) have described (Dean et al., 2014) and investigated (Dean et al., 2012) a suite of human-unique socio-cognitive abilities (also termed high-fidelity social learning mechanisms or social transmission mechanisms) posited to equip humans with the ability to transmit cultural traits with higher fidelity than is achieved by nonhumans. These high-fidelity transmission mechanisms include imitation (bodily action copying), teaching, pro-social behaviour and complex communication (language). Such mechanisms have been considered by numerous researchers (Boyd & Richerson, 1985, 1996; Tomasello, 1999; Legare & Nielsen, 2015; Lewis & Laland, 2012; Tennie et al., 2009; Tomasello et al., 1993, 1994; Whiten et al., 2009) to be candidate abilities underlying human cumulative culture because they enable cultural traits to be transmitted between individuals with high levels of precision. This has been suggested to lead to the persistence of traits within human cultural groups (as shown by mathematical modelling) and thus to allow more time for such traits to be improved upon by modification or recombination (Enquist et al., 2010; Lewis & Laland, 2012).

Proposals regarding high fidelity transmission mechanisms are, however, weakened by evidence that nonhumans can use such mechanisms (e.g. imitation in chimpanzees, Whiten et al., 1996, and teaching in meerkats, Thornton & McAuliffe, 2006) and yet do not display examples of cumulative culture (Galef, 2013; Heyes, 2012). Moreover, such mechanisms may not actually be necessary for cumulative culture (at least not in all circumstances): experimental evidence (Caldwell & Millen, 2009; Zwirner and Thornton, 2015) has shown that adults can accumulate cultural products without relying on imitation, teaching or language abilities – instead using information about end products (emulation learning). Dean et al. (2012) claim that these transmission mechanisms underlie human cumulative culture because during the solving of a novel foraging task they were either utilised only by their sample of human children (in the case of teaching), and not used at all in the comparison groups of chimpanzees and capuchin monkeys, or they were utilised substantially more by the children than the nonhumans (imitation and prosociality). However, although this correlational design provides some evidence that such abilities can be beneficial for problem solving, it does not show that these are pre-requisite requirements for cumulative culture (Galef, 2013; Heyes, 2012). The study instead suggests that humans are particularly adept at using high fidelity social learning mechanisms, and that they were more likely to succeed at this task, but the children might have found ways to solve the task, and

improve on others' solutions, even if they had been restricted from using these particular mechanisms (Reindl et al., 2020).

Many researchers now acknowledge that social learning mechanisms alone do not account for the human propensity to accumulate culture (Galef, 2013). Social learning mechanisms may ensure the continued presence of cultural traits within populations, but the definition of cumulative culture necessitates that traits increase in functionality. This is likely to occur through copying errors, novel invention, modification and/or recombination of cultural traits all of which fall under the umbrella term innovation (Enquist et al., 2008; Henrich et al., 2008; Legare & Nielsen, 2015; Lewis & Laland, 2012).

I have discussed issues with the proposal that social learning mechanisms provide us with an answer regarding the uniqueness of human culture (e.g. Dean et al., 2014). However, I now return briefly to the idea that social cognitive abilities, specifically our understanding of others' minds, is important for cumulative culture (Herrmann et al., 2007; Tomasello, 1999). Explicitly understanding the mental state of others (i.e. possessing a theory of mind: Call & Tomasello, 1999; Doherty, 2009; Premack & Woodruff, 1978; Wimmer & Perner, 1983) may allow one to utilise social information more flexibly, enabling social learning to be directed towards the most effective models and aiding information ratcheting (Dunstone & Caldwell, 2018).

Social Learning Strategies and Explicitly Metacognitive Social Learning Strategies

When discussing social learning we refer to situations in which an animal's learning of a response is influenced by the presence of a similar response in another individual, whether through observation of, or exposure to the products of, the other individual's activity, or through interaction between the two individuals (Heyes, 1994; Heyes, 2012; O'Sullivan, 2015). Learning of this kind has been documented in a wide range of taxa – from fruit flies (Sarin & Dukas, 2009), to chimpanzees (Whiten et al., 1999) and humans (Wood et al., 2013). Although using social information can be hugely beneficial (as exemplified in human cumulative culture), theoretical models (Boyd & Richerson, 1985; Laland, 2004; Rendell et al., 2010) have shown that indiscriminate use is not; social information is most valuable if it is used flexibly and in conjunction with individual learning. That is, animals are expected to switch between using individual and social learning and to utilise social information in a biased manner in order to enhance fitness. This idea of selective social learning was first described by theoretical models which predicted the scenarios under which learning from others would be beneficial (Laland, 2004). Empirical work now supports these predictions and adults, children and a plethora of nonhuman species evidence the use of “social learning strategies” (SLSs), otherwise termed “transmission biases” (Heyes, 2016; Kendal et

al., 2018). SLSs can be considered rules according to which the behaviour of humans and nonhumans has been found to conform and can largely be divided into three types: what information should be copied, when one should copy, and who should be copied (Heyes, 2016; Kendal et al., 2018; Morgan et al., 2012). All three types describe “strategies” which enable more useful traits to be retained within populations and it has therefore been suggested that SLSs may be important for cumulative culture. For example, a bias to copy the information which gives the highest payoff (a “what” SLS), to copy when personally uncertain (a “when” SLS) or to learn from the most successful individual (a “who” SLS) bolsters the chance that a more successful trait variant will be added to a population. However, there is a problem with this theory: if SLSs are responsible for cumulative culture we would expect them to be unique to humans, or at least not quite so prevalent in nonhumans. It has thus been proposed that human SLSs may be fundamentally different to the nonhuman variety (Heyes, 2016).

The word “strategy”, as in social learning *strategies*, implies that these are used with awareness and deliberate intent, yet this is misleading because the underlying cognitive mechanisms have not been determined. A strategy is ordinarily considered to constitute a plan to achieve a long-term goal but most SLSs (in both humans and nonhumans) are likely applied by default and reliant on associative learning mechanisms (Heyes, 2016; Kendal et al., 2018). In contrast, it has been suggested that some human SLSs, dubbed metacognitive social learning strategies, may be fundamentally different. These human-unique SLSs may rely on higher-level, System 2 or explicit cognitive, as opposed to associative, processes (Heyes, 2016; Kendal et al., 2018) and therefore more adequately warrant the label of strategy. SLSs which rely on associative learning processes have been termed “planetary” by Heyes (2016, 2018) because these “strategies” are not represented in the minds of the those whose behaviour conforms to them. In this sense the “users” of such SLSs can be compared to planets as the laws to which their behaviour conforms exist only in the mind of the observer. In contrast, metacognitive SLSs have been labelled “cook-like” because a cook following a recipe is aware of the rules they are adhering to when cooking as an individual using a metacognitive SLS is said to be aware of its content (and therefore have the ability to report such content). Heyes (2016) proposes that metacognitive SLSs enable a user to explicitly represent a rule regarding who, what or when they should copy and that this enables more selective use of social information, leading to ratcheting. However, it is slightly unclear as to why internal representation and use of a rule would provide an advantage over implicit representation resulting in use of the same rule. Heyes (2016) suggests that metacognitive SLSs provide more specific, precise rules which enable more knowledgeable individuals to be targeted. According to Heyes, planetary SLSs are likely to rely on

associative learning processes which have evolved through genetic evolution. Although associative learning allows such SLSs to be updated in line with an individual's recent experience, this is only useful up to a point because one individual's experience cannot capitalise on the experiences of an entire population. Contrastingly, cook-like, metacognitive SLSs evolve *through* cultural evolution *for* cultural evolution – they are learned socially and thus encompass experience gained by many individuals, over a greater time period, in different environments. This is said to make metacognitive SLSs (e.g. “copy the majority only when payoffs are visible”) highly accurate in directing an individual towards the best available information. Moreover, it is hypothesised that this precise quality of metacognitive SLSs means that only a small number of – carefully chosen – individuals will be copied, which is less risky than indiscriminate social learning. Furthermore, appropriate model choice makes blind copying very effective, thus allowing for high fidelity transmission of beneficial traits, aiding cumulative culture (Heyes, 2016, 2018).

Related to the above proposals, it has been suggested (Dunstone & Caldwell, 2018; Heyes, 2016) that human-unique ability for explicit (System 2, Evans & Stanovich, 2013) metacognition (defined as monitoring aspects of one's own cognitive processes, Flavell, 1979; Kloo et al., 2017) may allow more selective use of social information. Explicit metacognition relies on slow, conscious and rule-based processes (which tax executive functions) as opposed to implicit processes which are faster, non-conscious and automatic (and do not tax executive functions). Here explicit metacognition refers to understanding of others' minds (mentalising or theory of mind) as well as one's own (introspection). That is, it refers to a general cognitive capacity for explicit metarepresentation. Both mentalising and introspection have been proposed to enable more beneficial social information use and ratcheting. Dunstone and Caldwell (2018) describe various situations in which these self and other representations may optimise social information use due to changes to the behaviour of social information receivers and senders, or both. Common amongst these is an ability to actively change the way one interacts with social information due to an explicitly metacognitive understanding. For example, an ability to explicitly represent another's knowledge state (a component of theory of mind, thought to be late developing and human-unique, Call & Tomasello, 1999; Wellman et al., 2001) might enable one to recognise, and selectively copy from, a more knowledgeable individual based on this understanding (overriding any more general-purpose, default SLS). Although Heyes (2016) has proposed that most metacognitive social learning strategies are culturally transmitted, a strategy such as this might not even need to be learnt through social interaction but could be actively devised in a specific situation based on explicitly metacognitive reasoning. The individual using the strategy would therefore not only know *that* they are using it, and be able to report it, but

would be expected to understand *why* they were using it. I return to this idea later in the introduction (pp. 30-33) and investigate it in Chapters 4 and 5. Previous theories have cited theory of mind as a possible factor in explaining cumulative culture (see pp. 12-14), but the focus has been on how this might allow interpretation of others' actions, rather than the role that it might play in the selection of appropriate models.

The idea that explicit understanding of others' minds may enable one to engage in more selective social learning through actively devising an optimal social learning strategy bears some similarities with the technical reasoning hypothesis recently posited by Osiurak & Reynaud (2020) and discussed below. It is suggested that technical reasoning enables humans to evaluate the relevance of information provided by others and copy that which is most relevant. This seems in some way akin to recognising and selectively copying a more knowledgeable individual due to explicitly metacognitive understanding. Both theories emphasise the likely role of reasoned understanding based on cognitive abilities that are not assumed to be specialised for the acquisition of the contents of culture.

Physical Intelligence and the Technical Reasoning Hypotheses

In opposition to the aforementioned cultural intelligence hypothesis (Herrmann et al., 2007), Osiurak et al. (2016) recently posited the physical intelligence hypothesis: cumulative culture is limited by physical cognition. Physical cognition, in this respect, encompasses aspects of both fluid and crystallized intelligence – reasoning about the properties of a tool or technology in order to make mechanical inferences (technical reasoning, De Oliveira et al., 2019; Osiurak et al., 2020; Osiurak & Reynaud, 2020) and knowledge gained through previous, relevant tool-use. Osiurak et al. (2016) acknowledge that the evolved socio-cognitive skills described by the cultural intelligence hypothesis may enable faithful reproduction of a behaviour but believe that they do not explain the improvements to cultural traits observed in cumulative culture (at least not in the technological domain). According to their hypothesis, the uniqueness of human cumulative culture may be the result of humans possessing better physical cognitive abilities than nonhumans, rather than advanced social cognition. As previously discussed, research demonstrating cumulative culture in the absence of high-fidelity social learning mechanisms, such as imitation and teaching (e.g. Caldwell & Millen, 2009), provides evidence that socio-cognitive abilities are likely not essential for all cases of cumulative culture. Osiurak et al. (2016) utilised Caldwell and Millen's (2009) micro-society task and found that learners' physical intelligence (measurements of physical reasoning and visuospatial construction) predicted cumulative improvement to paper planes (distance travelled when thrown), but theory of mind did not. Furthermore, De Oliveira et al. (2019) reported similar findings but with a different task –

human-unique technical reasoning, but not theory of mind, predicted improvements to tower heights. However, both studies (De Oliveira et al., 2019; Osiurak et al., 2016) employed correlational designs, unlike Caldwell and Millen (2009) who manipulated the task design to prevent use of the capacities of interest (imitation and teaching). Therefore, these studies fall short of showing that the ratcheting observed was dependent on technical reasoning, as they only show that individuals with better technical reasoning abilities performed more highly.

When considering how better physical intelligence may enable the improvements to cultural traits observed in cumulative culture, Osiurak et al. (2016) relate their work to that on social learning strategies (pp. 14-17). They postulate that this type of intelligence could allow for more selective use of social information through an ability to evaluate the value of that provided by one or multiple individuals, resulting in copying of the most relevant information and increasing efficiency. This idea is extended in Osiurak and Reynaud's (2020) technical reasoning hypothesis in which it is proposed that human-unique technical reasoning skills may provide the foundation of cumulative culture through underlying both imitation and innovation. That is, technical reasoning may enable humans to acquire and copy the most relevant/best variant of a cultural trait via social learning, extracting relevant information and discounting that which is irrelevant through reasoning. Additionally, humans may improve extracted cultural traits via individual learning, using reasoned trial-and-error strategies and testing these to produce more efficient, functional products.

Demography

Demographic changes are frequently cited to have impacted the prevalence of cumulative culture across human societies (Dean et al., 2014; Henrich, 2004; Richerson, 2013) and mathematical models have shown that larger numbers of interacting individuals (an "effective" population) can result in cumulative improvements (such as during the upper Palaeolithic 45000 years ago, Powell et al., 2009; Shennan, 2001) or protect against loss of cultural complexity (see "The Tasmanian Case", Henrich, 2004). If one assumes that learners within a population can identify the most appropriate individual to copy (see pp. 14-17), and that copying is imperfect, then larger populations increase the likelihood that any individuals copied hold traits of higher than average fitness. Therefore, at some point, the benefit of a larger population will outweigh the losses from imperfect copying and traits can accumulate beneficial modifications (Henrich, 2004; Kobayashi & Aoki, 2012). Support for this association between population size and cumulative culture is provided by ethnological studies (e.g. Collard et al., 2013; Kline & Boyd, 2010) and an emerging body of experimental work (Dereux et al., 2013; Kempe & Mesoudi, 2014; Muthukrishna et al., 2014) which

demonstrates that increasing the number of social models (from one to multiple models) can result in ratcheting. However, it appears that under certain circumstances, more models, and thus more information, may not be beneficial for improving cultural traits, and may even be detrimental (Caldwell & Millen, 2010; Fay et al., 2019). I discuss challenges associated with utilising information from multiple models in Chapter 3.

The aforementioned mathematical models assume that the cognitive capacities required for ratcheting were in place prior to any effects of demographic factors (as noted by Dean et al., 2014; Henrich, 2004) therefore demography alone cannot explain the human-unique propensity for cumulative culture. Nevertheless, the fact that demographic factors exert such influence on human cultural traits may enable insight into the cognitive capacities required to exploit them. In Chapter 3, I thus investigate how ability to utilise information from multiple models changes across development.

Brief Review of Experimental Methods Used to Investigate Cumulative Culture

I have discussed relevant theoretical background on the study of cumulative culture and I now move on to briefly consider the main methods used by scientists in their attempts to examine the phenomenon in humans (including human children) and nonhumans. I later outline the methods I have used to investigate cumulative culture and selective social learning in children (see pp. 30-33).

Cumulative culture has been investigated using theoretical models, experimental methods (both in and outside of the laboratory), and field studies of natural populations (Caldwell & Millen, 2008a, 2008b). Experimental methods, as used in this thesis, can be particularly beneficial in elucidating the abilities which may underlie cumulative culture because they provide scope to manipulate key variables of interest and discover any prerequisites (cognitive or behavioural) for ratcheting under different experimental conditions (Caldwell, 2018). I therefore focus on offering a brief review of such methods. Theoretical models and field studies will not be reviewed as these methods are outside the scope of the current thesis.

Generational Replacement Methods

In order to provide evidence of cumulative culture, it is important to identify whether learner benefits accrue over generations such that a learner achieves greater benefit by copying an individual from a later, rather than an earlier, generation i.e. whether the social information in later generations is more valuable. The generational replacement approach (encompassing the replacement micro-society and transmission/diffusion chain methods – see the top image in Figure 1.1) allows capture of this important aspect through simulation of

generational succession within a group of participants. This approach has transformed the study of cumulative culture in adult populations (e.g. Caldwell & Millen, 2009; Morgan et al., 2015; Muthukrishna et al., 2014; Zwirner & Thornton, 2015), enabling this group-level process to be modelled using a relatively small number of individuals within a controlled environment. The method has also been utilised by researchers examining cumulative culture in children (Flynn & Whiten, 2008; Reindl & Tennie, 2018; Tennie et al., 2014) and nonhumans (Sasaki & Biro, 2017), although both bring logistical challenges.

Caldwell and Millen (2008a, 2008b, 2009) pioneered use of the replacement microsocociety approach (originally used by Jacobs & Campbell, 1961, for an alternative purpose) to model cumulative culture in the laboratory. Small populations of learners (miniaturised societies) are created and members of the population are continuously removed and replaced with naive members in order to simulate the required generational succession. Each learner is given the same task, under the same conditions, with a defined goal (e.g. 5 minutes to build a tower as tall as possible (Caldwell & Millen, 2008a), and all learners from a particular generation can interact with those from a number of previous generations. The ability to define a specific end goal is important because it enables researchers to evaluate whether performance improves over generations, and thus whether information provided by later generations is more valuable, in relation to this set goal, when all other variables remain constant. Note that generational improvements in performance can encompass changes to functionality and/or efficiency, depending on the task goal (see Caldwell, 2018; Wilks & Blakey, 2018, for further discussion of this distinction). A similar approach, also utilising generational replacement and used to study cumulative culture, is termed the transmission/diffusion chain (Mesoudi & Whiten, 2008; also see the top image in Figure 1.1). There are different variations of this method, but it primarily differs from the replacement microsocociety in three main ways: each participant obtains information from just one cultural parent, information transfer is unidirectional (vertical, from cultural parent to cultural offspring), and members of different generations do not perform the same role simultaneously (e.g. there would not be members teaching in both Generations 3 and 4 at the same time).

Following the successful use of generational replacement methods to evidence cumulative culture in an adult population (e.g. Caldwell & Millen, 2008a), these methods have now been used to investigate the capacity for cumulative culture in children (Reindl & Tennie, 2018; Tennie et al., 2014) and nonhumans (Sasaki & Biro, 2017). However, generational replacement methods can also be used to investigate the requirements for ratcheting under different task contexts (e.g. see Caldwell & Millen, 2009, 2010; Fay et al., 2019; Zwirner & Thornton, 2015). Experimental manipulations, such as limiting use of a

particular social learning mechanism (Caldwell & Millen, 2009; Zwirner & Thornton, 2015), have enabled researchers to identify, within the context of a specific task, the constraints on ratcheting and hence any cognitive or behavioural prerequisites. Theories which posit that the uniqueness of human cumulative culture relies on certain abilities (e.g. particular social learning mechanisms – see pages 12-14) can therefore be tested by excluding use of the variable of interest or manipulating this across a number of experimental conditions. For example, Caldwell and Millen (2009) found that cumulative improvements to paper aeroplanes, as found in Caldwell and Millen (2008a), persisted when imitation and teaching were prevented by an experimental manipulation. We return to examine task context in Chapter 2, although we use an alternative experimental method, as described in “An Alternative Method” and depicted in the below image in Figure 1.1.

Based on the points outlined above, it is clear that generational replacement designs (Caldwell et al., 2020) have significant advantages. However, in practise implementation may be challenging due to the logistics of coordinating participants, and/or the information which needs to be transferred between them (often via personal interaction), in space and time. This can be particularly difficult when performing research with nonhumans (see Menzel et al., 1972) or young children who cannot easily be directed to comply with strict experimental procedures. Furthermore, these methods require large numbers of participants – a major limitation when investigating restricted populations such as endangered nonhuman species or children in multiple, narrowly defined age bands, as is necessary for investigating how capacity for ratcheting may change across development (a major component of this thesis, Chapters 2 and 3). As stated above, generational replacement designs do allow investigation of ratcheting under different restrictions and contexts, but this requires even larger numbers of participants to allow for multiple conditions. Unfortunately, such important research questions therefore compound the issues of sample size when working with nonhumans and young children. We address this issue in Chapters 2 and 3 with a recently proposed method (see “An Alternative Method”, Figure 1.1 and Caldwell et al., 2020) particularly suited to developmental research into age and context-related effects on ratcheting.

Closed Group

Closed group methods have been utilised to investigate cumulative culture when generational replacement methods (pp. 19-21) would be logistically challenging and impractical. For example, when working with children (Dean et al., 2014; McGuigan et al., 2017) or nonhuman (Dean et al., 2012) populations. However, the conclusions which can be drawn from these methods are more limited. In contrast to generational replacement studies,

those utilising closed groups encompass introducing a problem, usually some kind of apparatus, into a group (population) of participants with fixed or “closed” membership. The problem contains a range of potential solutions which are deemed to vary in difficulty, with higher-level solutions supposedly building on learning required to access lower level solutions (although this is not always clear, see our discussion of Dean et al., 2012, pp 25-26). Participants are permitted to interact freely, engaging in numerous task attempts, so there is the potential to investigate whether learning of the higher-level solutions is built up in a cumulative fashion through multiple rounds of social learning. For example, a group member discovers a solution which other group members learn through social learning and this aids another group member to discover a higher-level solution which also spreads via social learning. However, although these studies appear to have the potential to identify cumulative culture, in practise it is difficult to differentiate whether the higher-level solutions are discovered by individual learning such as trial-and-error (perhaps aided by the social environment), or social learning i.e. whether learning these solutions socially would confer any benefit to new learners entering the group over and above individual learning.

An Alternative Method

I have reviewed generational replacement and closed group methods, which have, until very recently, represented the two best options for measuring cumulative culture experimentally. However, as alluded to above, there are some issues with the use of these methods for the investigation of cumulative culture in children and nonhumans. A major failing of the closed group method is that it is impossible to determine whether new learners entering the group would learn higher-level solutions due to benefit from accrued social information, or rather, as a result of individual learning (i.e. the third criteria for cumulative culture according to Mesoudi & Thornton's, 2018, specification outlined on page 9, “improvement in performance as a consequence of the modification”). The generational replacement method does not suffer from this issue, allowing the performance of learners in earlier and later generations to be directly compared whilst all other task variables remain constant, but it does present logistical problems and require a large number of participants.

A relatively new approach (introduced by Caldwell et al., 2020) renders it possible to retain the advantages of the generational replacement method and yet eliminate the associated sample size and logistical problems which limit its implementation for the study of nonhumans and age-related changes in children. We utilise this method in Chapters 2 and 3. Rather than measuring task performance in many individuals, each completing a small number of trials (as in generational replacement methods), this method requires a much smaller number of participants each of whom complete multiple trials. The new method thus

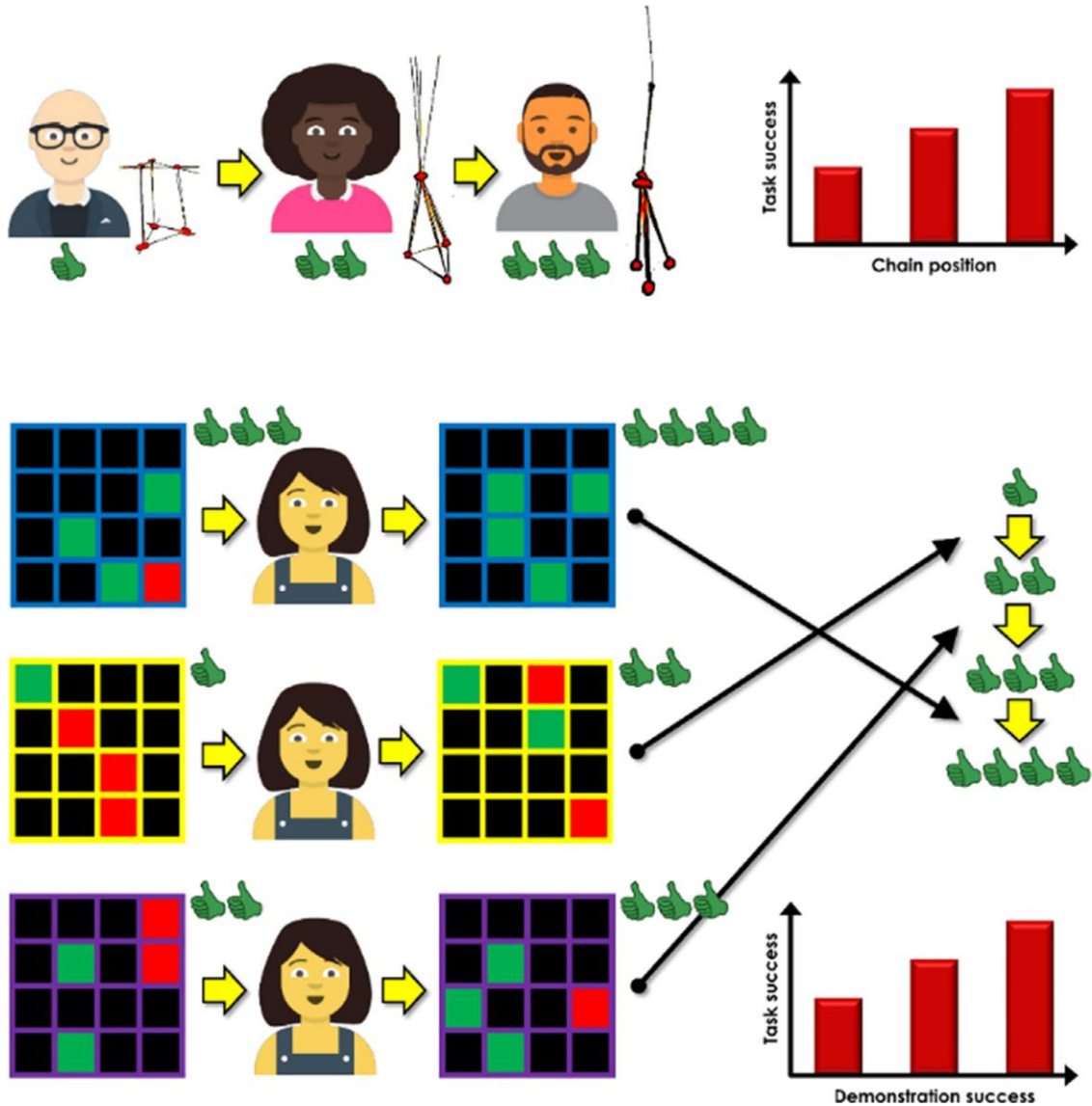
allows the outcome of repeated transmission episodes – the potential for ratcheting (cumulative culture) – to be determined using data from individual participants.

Cumulative culture is a group level phenomenon therefore the most effective methods of investigating this have typically involved testing participants in groups. Consider the generational replacement methods – we can determine whether learner benefits accrue over generations through measuring whether learners in later generations perform better at the task than those in earlier generations. Although testing individual participants may seem counterintuitive, the new method allows measurement of a ratchet effect in each participant, negating the necessity of physically forming a chain (Figure 1.1). In this method the outcomes from a number of different generations (improvement in task performance) within a transmission chain can be simulated using the data from individual participants: each participant is exposed to a number of demonstration types differing in value (e.g. containing more or less rewards – see the below image in Figure 1.1 and Chapters 2 and 3), enabling us to investigate whether higher value demonstrations (akin to those in later generations of a transmission chain) result in better performance than those of lower value (akin to earlier generations in a transmission chain). That is, we can determine whether the participant is able to exploit the information available from each demonstration and perform more highly following better-quality social information. We can thus determine whether ratcheting would occur in a chain of participants in which the information they received was comparable to that presented to individual participants in this task (see both images in Figure 1.1).

When implementing the new method it is important that the task in which ratcheting is measured conforms to certain restrictions. Most tasks which have been used to investigate cumulative culture involve presenting the same problem, with a fixed solution space (e.g. recreating a target image, Muthukrishna et al., 2014, or making a paper aeroplane, Caldwell & Millen, 2008a) to each generation of participants whose individual solutions will differ in quality. However, the new method involves presenting each participant with multiple solutions (demonstrations) of different value (again, see below image in Figure 1.1), representing different points in a theoretical transmission chain, and determining the effect of each solution on participant performance at this point in the theoretical chain. Presenting the same problem over multiple demonstrations would prevent accurate measurement of cumulative culture because a participant would not be naïve to each of the different demonstrations, preventing their use as representations of different positions in a transmission chain. Instead, the problem presented in each generation needs to be solved independently (as in a transmission chain) and yet be of the same type as the problems in the other generations. The task presented is thus likely to be relatively abstract. For example (as utilised in Chapters 2 and 3, and shown in Figure 1.1), the type of problem could be

finding a set number of rewards hidden within arrays of a set size, with the different solutions constituting information about the reward locations for individual arrays, each with their own, independent reward distribution.

Figure 1.1

An Alternative Method for Assessing Cumulative Culture

Note. Top image: solution success increases in a chain of multiple individuals attempting a single problem. Below image: solution success increases with demonstration success in single individuals attempting multiple problems. Adapted from “Experimental assessment of capacities for cumulative culture: Review and evaluation of methods,” by C.A. Caldwell., M. Atkinson., K. Blakey., J. Dunstone., D. Kean., G. Mackintosh., E. Renner., and C.E.H. Wilks, 2020, *Wiley Interdisciplinary Reviews: Cognitive Science*, 11(1), p. 1. Copyright 2019 The Authors. Adapted with permission.

The practical benefits of the new method, in relation to the investigation of cumulative culture in children and nonhumans, have been stated above. Importantly, this also enables direct comparisons between children, adults and nonhumans (completing the same, or a very similar, task) and thus presents an exciting opportunity to investigate any constraints on ratcheting, and associated cognitive mechanisms, which could account for differences in cumulative culture between humans and nonhumans. Of particular relevance to this thesis is the potential to use narrowly defined age bands to study children's ability to ratchet information at different stages of cognitive development, and in different contexts (see pp. 30-33 and Chapters 2 and 3).

Although useful for addressing my research questions in children, the method described above is not without limitations. In particular, the abstract nature of tasks which utilise the new method may not be ecologically valid enough to address those of researchers interested in more "real-world" social information use. Furthermore, this new approach provides limited scope to address research questions on group size and structure. The effects of population size can be studied by increasing the amount of information presented to a participant (see Chapter 3) but investigation of bidirectional communication (such as interactions between group members) is not possible. Whilst the new method enables measurement of ratcheting across humans and nonhumans, use in the latter is likely to require substantial training, rendering this method less appropriate for those investigating spontaneous social information use (Caldwell et al., 2020).

Studies of Cumulative Culture in Children

Adult humans frequently display examples of cumulative culture, exemplified experimentally using the generational replacement designs described on pages 19-21 (e.g. Caldwell & Millen, 2008a, 2009, 2010; De Oliveira et al., 2019; Fay et al., 2019; Kempe & Mesoudi, 2014; Morgan et al., 2015; Muthukrishna et al., 2014; Osieurak et al., 2016, 2020; Zwirner & Thornton, 2015). Therefore, we can make the logical assumption that the capacity for cumulative culture must present at some point during a child's development. This means that children are an ideal group in which to investigate why cumulative culture becomes possible and, furthermore, why it is so widespread across human societies but appears to be extremely limited (or absent) in nonhumans. More specifically, the development of human-unique cognitive capacities (including some of those discussed on pages 12-18) may predict shifts in children's use of social information and thus allow us to elucidate which key cognitive capacities may support ratcheting (note – this may differ with context, see pages 30-33). Although it is unlikely that very young children are capable of ratchet-like cultural transmission, because they are not yet equipped with adult-level social learning skills, the

cognitive and behavioural abilities required for ratcheting are expected to emerge with age due to brain maturation and experience. Performing experimental work to investigate ratcheting in children across a wide age range therefore offers the exciting potential for determining when shifts in social information use occur and any concomitant cognitive advances. Such experimental work, and the logic outlined above, underpins this thesis (as further detailed on pages 30-33).

Despite the potential knowledge gains stated above, designing suitable experiments to detect cumulative culture in children is challenging and measuring age related changes adds further difficulties. For example, utilising the highly successful generational replacement method (Caldwell & Millen, 2008a) to measure age-dependent changes would be logistically challenging and require large numbers of participants, in narrowly defined age bands. This is likely to be the reason that only a small number of researchers have attempted to test for cumulative culture in children and have done so using single age groupings. In theory this approach allows investigation of the presence or absence of ratcheting, in a particular context and age group, but it excludes study of potentially informative, age-related changes. Nevertheless, some research, utilising either closed group (Dean et al., 2012; McGuigan et al., 2017) or generational replacement (Reindl & Tennie, 2018; Tennie et al., 2014) designs, is available for perusal. We now give an overview of this work and consider the implications.

Closed group designs offer the potential to investigate cumulative culture in children and nonhumans when the use of generational replacement designs is difficult or impractical (pp. 21-22). I have previously outlined a study of this type, conducted by Dean et al. (2012), in my discussion of whether social learning mechanisms underlie cumulative culture (pp. 12-14). However, I now return to this study, and a similar one conducted by McGuigan et al. (2017), to consider the contribution to our understanding of children's capacity for cumulative culture.

Dean et al. (2012) investigated cumulative culture in children (aged 3 to 4) and nonhuman primates (capuchin monkeys and chimpanzees) using the same task. A puzzle box which could be solved via three solutions varying in difficulty, designed to represent accumulated behaviours, was placed into social groups of each population. The solution designed to be the easiest was to push a door, and this gained participants a low-grade reward. Pressing a button, and sliding the door further, was deemed to be more difficult and thus resulted in a medium-grade reward. Finally, twisting a dial to move the door further still was considered the most difficult solution and revealed a high-grade reward. Children found both the medium and most difficult solutions much more often than non-human primates; in fact, most primates did not progress past obtaining the lowest grade reward. Although

children learnt behaviours designed to build on simpler ones, this does not enable us to conclude that they showed cumulative culture because it was not possible to differentiate the impact of individual learning based on task experience from that of exposure to social information i.e. whether learning task solutions from others would confer any benefit to new group members over and above individual learning such as trial and error.

Children alone showed high propensity for social learning, with the use of sociocognitive capacities positively correlated with the level of task solution reached. This led the authors to conclude that nonhumans do not have the capacity for cumulative culture and that specialised sociocognitive processes, absent or diminished in nonhumans, underlie human cumulative culture (Dean et al., 2012, 2014). However, as previously discussed (pp. 12-14), the correlation between children's use of sociocognitive processes and reaching of higher-level task solutions does not prove that such processes are essential for ratcheting. To show that a behaviour, or aspect of cognition, is a pre-requisite (for a particular population and context) would require proof that ratcheting can proceed in its presence but not in its absence.

McGuigan et al. (2017) studied cumulative culture using a similar design, also inserting a puzzle box into groups of 3-4-year-old children and permitting free interaction. As in Dean et al. (2012), this study involved solving problems to reach rewards of varying value. There were four levels of accumulating difficulty (associated with rewards of increasing desirability), designed to enable children to build upon solutions discovered on previous levels and thus allow determination of their potential for cumulative culture. For a Level 1 reward children needed to move a capsule containing the reward through one of four exits using their fingers. Level 2 involved a similar set-up, but a tool needed to be used (the idea being that knowledge gained through using fingers to manipulate rewards into the correct position at Level 1 could be applied to tool use at Level 2). At Level 3 children were required to attach two tools together to make a longer tool (or unfold a folded tool), and at Level 4 needed to fashion a hook. Furthermore, the four exits at each level were designed to allow children to generate, and build-upon, solutions at previous levels. The exits were of the same type across levels, but each exit was opened in a different way, therefore the authors postulated that knowledge of an exit at a lower level could be re-used at a higher level, albeit with a more complex form of tool.

Children frequently gained success at Level 1, with success at the more challenging, second and third, levels accumulating as the experiment progressed. Children also engaged in social learning, e.g. they were more likely to release a reward through an exit which they had previously seen used by another child, and the success at higher levels was not replicated in an asocial control condition. It was thus concluded that there was sufficient

evidence of a developing capacity for cumulative culture in young children. Yet, as acknowledged by the authors themselves, the claims regarding social transmission should be considered with caution because these were not concluded from direct experimental manipulation (as is the case when using generational replacement designs). To ascertain that social learning was necessary and sufficient to achieve the higher-level solutions, one would need the ability to separate the benefits gained as a result of exposure to social information from those resulting from an individual's own task experience. As in Dean et al. (2012), such separation was not possible using this design and, although the higher-level solutions may have been learnt socially, one cannot conclude whether specifically learning from later generations would confer any performance benefit over and above one's own increasing experience with the task and/or the potentially bolstering effect of learning within a social environment (e.g. having one's attention drawn to a particular part of the experimental apparatus by another child). In summary, this work does not conclusively demonstrate cumulative culture, although the results do suggest that groups of children aged 3-4 may have the ability to ratchet easily accessible information.

Following the success of generational replacement designs for demonstrating cumulative culture in adult humans, two studies (Reindl & Tennie, 2018; Tennie et al., 2014) have attempted to use this method to test for a ratchet effect in young children. Unlike closed group designs, such experiments enable social transmission of a modified trait to be tracked over generations of learners (see pp. 19-21) – a crucial requirement for cumulative culture (Mesoudi & Thornton, 2018).

Tennie et al. (2014) examined whether children (aged 4) would show a ratchet effect on a task in which they were required to transport as much rice as possible from one location to another using a range of tools provided. Children were organised into transmission chains ($n = 5$) and thus had the opportunity to observe the previous child's transportation attempt prior to selecting a tool and attempting the task themselves. No evidence for a ratchet effect was found and transportation solutions in the final generations were no more efficient (i.e. did not lead to the transportation of more rice) than those spontaneously adopted by children in the first generations who had had no exposure to social information. In a second experiment, the experimenter acted as the first generation in the chains and demonstrated an inefficient transportation method – using an inappropriate tool chosen from the same set presented to children in the first experiment. In this case, half of the chains underwent a switch to the use of a more efficient method, using a more suitable tool. This method was transmitted along the chains and resulted in those in the final generations transporting more rice.

Similarly, Reindl and Tennie (2018) investigated children's potential for cumulative

culture using transmission chains of ten generations. Children aged 4 and 5 were given an adapted version of Caldwell and Millen's (2008a) tower-building task in which the goal was to build as tall a tower as possible from sticks and plasticine. In order to minimise logistical difficulties, children did not physically interact but were presented with towers built by children from the previous two generations, or a replica of these (except in Generations 1 and 2 who had access to zero and one previous child's attempts respectively). Although there was evidence for social transmission, tower height did not increase over generations of learners. Therefore, as in Tennie et al.'s (2014) first experiment, there was no evidence of a ratchet effect.

The study by Reindl and Tennie (2018), and the first experiment in that by Tennie et al. (2014), could be considered to show that cumulative culture is beyond the capabilities of young children, at least for the behaviours measured in these tasks. However, Tennie et al.'s (2014) second study (and also Flynn, 2008, see Chapter 2) provides evidence that children were able to improve upon methods which were *less* effective than those developed spontaneously in the absence of any social information. Such improvements would not be considered an example of cumulative culture by most researchers in the field, and some have even coined the term “subtractive ratchet effect” to differentiate these (e.g. Tennie et al., 2014). Nonetheless, these findings suggest that under certain circumstances ratcheting may be possible even in very young children and thus highlight the need for further research in this area.

At present we know very little regarding the circumstances under which ratcheting may, or may not, be possible in young children and the underlying cognitive capacities required. Collectively, the studies referred to above (Dean et al., 2012; McGuigan et al., 2017; Reindl & Tennie, 2018; Tennie et al., 2014) tentatively suggest that this process might occur from around three years of age. However, these experiments focus on determining whether children of one particular age do, or do not, show cumulative culture in one particular circumstance. Yet, young children (and possibly nonhumans) may be able to ratchet information under a restricted range of circumstances compared to adult humans. We discuss this further in Chapter 2. Relatedly, restricting experiments to single age groupings has limited the ability of scientists to investigate age related changes in the capacity for cumulative culture. If a carefully designed task measured children's ability to show a ratchet effect across a range of ages, the pattern of task performance across development could provide an indication as to the cognitive capacities required. I elaborate on this point in the section below.

Contribution of This Thesis

Broadly speaking, this thesis investigates the cognitive mechanisms which may underlie human propensity for cumulative culture through studying children between 3 and 10 years old. Although it is accepted that human adults are capable of accumulating culture, experimental research in children (as reviewed on pages 26-30) is limited and it is yet to be determined if, and under what circumstances, children have the capacity for cumulative culture. Furthermore, studying children at different stages of cognitive development can help us to understand the constraints on ratcheting at different ages and elucidate which particular cognitive mechanisms may, or may not, underlie human propensity for cumulative culture. In regard to this, a number of experiments have been conducted in adults (discussed on pages 19-21), but at present research in children has focused on attempting to show whether or not children display ratcheting (pp. 26-30) rather than investigating the underlying cognitive mechanisms. The basic premise for the novel work presented in this thesis is thus that if children begin to use social information differently around the time that a particular human-unique cognitive ability develops then this indicates that the ability may be a fundamental requirement for cumulative culture, or at least in the particular context studied.

Candidate cognitive abilities may include those which represent a significant milestone during human development, an example of which is the capacity to understand the minds of others: theory of mind (Premack & Woodruff, 1978; Wimmer & Perner, 1983). This ability appears to develop rather suddenly in children between the ages of 3 and 4 (Doherty, 2009; Wellman et al., 2001; Wimmer & Perner, 1983) and, although this remains a topic of debate, it does not appear to be shared by nonhuman primates (Call & Tomasello, 2008; Heyes, 2015; Whiten, 2013). Nonhuman primates' ability to understand others' minds appears to be no more advanced than that of a human infant (Hare et al., 2000, 2001; Moll & Tomasello, 2006; O'Neill, 1996). If theory of mind, or related abilities such as understanding of others' knowledge (Pillow, 1989; Pratt & Bryant, 1990; Schmidt & Pyers, 2011; Wimmer et al., 1988), are implicated in cumulative culture then we might expect to see a sudden shift in social information use in line with the development of the relevant capacity. I consider theory of mind, and related abilities, in Chapters 4 and 5 (detailed further below).

It is also possible that cognitive abilities which develop more gradually throughout childhood may be implicated in human capacity for cumulative culture. In this case, the premise outlined above still applies as we may observe changes in social information use across a developmental range in a condition which taxes the ability implicated, compared with minimal or no changes in a less taxing condition. Working memory (defined as the storage and processing/manipulation of limited information, Best & Miller, 2010; Cowan, 2008; Diamond, 2013; Garon et al., 2008) could be considered an example of such a cognitive ability because it develops throughout childhood and into adolescence (Best &

Miller, 2010; Diamond, 2013; Garon et al., 2008). Although the small amount of research on nonhuman working memory suggests that some nonhuman mammals may have similar storage capacities to humans (Carruthers, 2013), humans make better use of working memory capacity through a range of strategies, the use of which increases with age (Armitage et al., 2020; Bulley et al., 2020; Carruthers, 2013). I consider working memory in Chapters 2 and 3.

The reviewed proposals for the uniqueness of human cumulative culture encompass aspects of social and physical cognition, pages 12-17 and 17-18 respectively. The beneficial effects of population size have also been outlined, and we have posited that a cognitive explanation may underlie such effects (pp. 18-19). The aim of the work in the ensuing chapters was thus to expand on these ideas, suggesting that differences in the cognitive mechanisms available to humans and nonhumans may underlie cumulative culture. This was explored through two main research strands.

Firstly, (in Chapters 2 and 3) I will describe investigations of realistic constraints on ratcheting social information under different, ecologically valid contexts and will consider the aspects of cognition which may be implicated. In these studies I did not set out to test one specific hypothesis as to the cognitive capacities underlying cumulative culture, but rather, to consider the age at which children are able to ratchet information in these simplified contexts and what this tells us about the cognitive capacities which may be implicated in human cumulative culture. However, as these studies included contexts in which useful information remains present within the environment, or is masked, we postulated that working memory may be involved. I have described proposals relating social (pages 12-17) and physical (pages 17-18) cognition to our potential for cumulative culture but the involvement of general cognitive abilities, such as working memory, has been somewhat neglected in the literature to date. In both Chapters 2 and 3 I utilised a number of dependent variables to measure children's ability to effectively utilise social information presented in different ecologically valid contexts (e.g. transient or enduring, or when the information is separated in time or space) and their potential to ratchet this information.

In Chapter 2 I measured children's ability to copy useful information, and discount non-useful information, to improve on the score achieved by a social demonstrator in a simple game. Chapter 3 also explored the use of information in different contexts, but these were related to the contexts expected if information is to be utilised from multiple different individuals. I therefore analysed children's ability to capitalise on information presented by multiple models. The cognitive constraints this may present have previously been overlooked and certain cognitive capacities may be required in order to exploit the increased information available. Again, I measured children's ability to copy and discount useful and

non-useful information respectively. I also measured their potential for ratcheting, which in this study represented the ability to improve on the best information present within a 2-model population. The study I present is the first to investigate the cognitive abilities which may underlie the use of information from multiple models across development. The studies in both Chapters 2 and 3 were also the first to utilise a relatively newly proposed method (the advantages of which are discussed on pages 22-26) which enabled measurement of the potential for ratcheting in children of different ages, and thus stages of cognitive development (children aged 3-6 and 5-10 in Chapters 2 and 3, respectively), across different contexts.

In Chapters 4 and 5 I will turn to the question of whether humans use social information more selectively due to the human-unique ability to explicitly represent the content of another's mind (e.g. theory of mind and related developments). Theory of mind has previously been proposed to enable cumulative culture by aiding the understanding of others' intentions and goals, leading to the use of social learning mechanisms such as imitation and teaching (Tomasello et al., 1993, 1994; Tomasello, 1999). However, unlike this earlier work, in Chapters 4 and 5 I examine how explicit theory of mind may facilitate selection of the most suitable models. Such a proposal instead relates to the recent hypothesis (described on pages 14-17, Dunstone & Caldwell, 2018; Heyes, 2016) that human-unique explicitly metacognitive social learning strategies enable cumulative culture through the selective copying of more effective individuals/trait variants. The studies included in this thesis are, to my knowledge, the first to test this hypothesis in a population of children.

If capacities for explicit metarepresentation (specifically, explicit theory of mind and related developments) do enable more selective social information use we would expect there to be an association between milestones in human development of these capacities and ability to use social information selectively. In Chapters 4 and 5 I thus tested whether there was a specific point in development in which children (aged 3-8) were able to use social information more strategically by measuring whether there was a shift in the way in which information from an informed demonstrator was used. The tasks performed relied upon the assumption that if a participant had an explicit understanding that one of a number of social demonstrators held additional task-relevant knowledge they would selectively copy this informed demonstrator. I therefore measured whether there was an age at which children specifically chose to look for hidden objects in the same (correct) location as such a demonstrator.

Chapter 2: Cognitive Prerequisites for Cumulative Culture are Context-Dependent: Children's Potential for Ratcheting Depends on Cue Longevity

Note. This chapter has been accepted for publication in its current form:

Wilks, C.E.H., Rafetseder, E., Renner, E., Atkinson, M., & Caldwell, C.A. (in press).

Cognitive Prerequisites for Cumulative Culture are Context-Dependent: Children's Potential for Ratcheting Depends on Cue Longevity. *J Exp Child Psychol*.

Abstract

Human cumulative culture has been suggested to depend on human-unique cognitive mechanisms, explaining its apparent absence in other species. We show that the potential for exhibiting cumulative culture depends on the cognitive abilities of the agents and the demands associated with using information generated by others' activity. 154 children aged 3-6 years played a searching game ("Find the Treasure"), taking their turn after a puppet demonstrator. The puppet's attempt revealed information about the contents of the locations searched, which could be exploited to target rewarded locations, and avoid unrewarded ones. Two conditions were presented, intended to capture realistic variation in the transience of the cues generated by another individual's activity. In one condition, the puppet's demonstration provided transient information – boxes were opened, seen to be rewarded or not, and then closed. In the other condition the puppet's chosen boxes remained partially open, providing an enduring visible cue as to whether that location was rewarded. Children undertook three trials of varying demonstration success, and we used patterns of performance to infer the potential for improvement over multiple generations of transmission. In the Enduring Cues condition, children's performance demonstrated the potential for cumulative culture. In contrast, in the Transient Information condition, only older children showed improved performances following higher success demonstrations and overall performance was not compatible with the possibility of improvements over generations of social transmission. We conclude that under certain conditions cumulative culture could occur in many species, but in a broader range of contexts in humans.

Introduction

Cumulative culture typically refers to a particular subcategory of cultural evolution, characterised by a directional pattern of change that results in “improvements” (Tennie et al., 2009) or increasingly “preferred” traits (Caldwell, 2018). However, a more precise definition which satisfies the diverse range of scientists interested in this phenomenon is yet to be agreed upon. Mesoudi and Thornton's (2018) recent sampling of published work outlined four core criteria for cumulative culture: i) a change to a behaviour or cultural product, ii) social transmission of the modified trait, iii) improvement in performance as a consequence of the modification, and iv) iteration of these steps resulting in ongoing improvement over time. These criteria were common across the entire sample and thus can be considered a consensus definition. Yet, they do not encompass additional requirements adhered to by some researchers – classified by Mesoudi and Thornton (2018) as “extended criteria”. For the purpose of adhering to a precise definition, and one which is most representative of that used in the field, we therefore refer to the aforementioned core criteria when using the term cumulative culture. It should however be noted that some other definitions of the concept are more restrictive (e.g. Enquist et al., 2011; Hunt & Gray, 2003; Pradhan et al., 2012; Reindl et al., 2020). The notion of constant improvement has also led to cumulative culture being referred to as the *ratchet effect* (Tennie et al., 2009; Tomasello, 1990). We use these terms interchangeably.

In contrast to its ubiquity in human populations, evidence of cumulative culture in nonhumans is strikingly scarce (e.g. Dean et al., 2014), with some authors proposing that it is unique to humans (e.g. Tomasello et al., 1993). Some have even argued that cumulative culture depends on cognitive mechanisms which themselves are proposed to be unique to humans (Dean et al., 2012; Tennie et al., 2009). However, recent nonhuman evidence from both experimental (Sasaki & Biro, 2017) and field research (Jesmer et al., 2018) appears consistent with Mesoudi and Thornton's (2018) core criteria. These findings in pigeons (Sasaki & Biro, 2017), and bighorn sheep and moose (Jesmer et al., 2018), therefore suggest that cumulative culture is not precluded in nonhumans but may simply be far more restricted in its expression compared with human examples.

In line with this view, experiments on cumulative culture in humans suggest that transmission requirements depend on the context and behaviour in question. For example, Caldwell and Millen (2009), and subsequently Zwirner and Thornton (2015), found that imitation (in the sense of bodily action copying) and teaching are not always required for improvements in performance over multiple transmission events in simple building tasks; exposure to completed products was sufficient. However, imitation may be required for

ratcheting during more “cognitively opaque” tasks (e.g. Wasielewski, 2014). There is also little dissent regarding the value of teaching (Morgan et al., 2015) for transmission of skills that are more complex and/or cognitively opaque (Csibra & Gergely, 2009). Consistent with this, Caldwell et al. (2018) found that teaching facilitated the transmission of effective knot-tying techniques for complex knots, whereas simple knots were transmitted equally effectively from exposure to end products alone. Furthermore, Osiurak et al. (2020) showed that a teacher’s theory-of-mind ability predicted cumulative performance, but only when the teacher did not have visible access to a learner’s actions. Work focussed on the constraints which may underlie the tool innovation abilities of individual children, rather than the cumulative result of a social transmission process, is also relevant here; Neldner et al. (2017) provide evidence that the affordance visibility of a tool (whether one can easily perceive how the tool can be used to achieve a desired action e.g. presence of a visible hook for hooking) may affect the likelihood that task-relevant innovations to the tool occur. The examples outlined above thus raise the possibility that, for a population sharing a common pattern of cognitive abilities and behavioural proclivities, transmission patterns consistent with cumulative culture may be possible for some behaviours but not others. In populations exhibiting a different suite of traits, the range of contexts in which cumulative culture might be manifested could be more extensive, or more limited.

We consider human children to be a particularly interesting group within which these ideas can be tested. Firstly, we know that human adults are capable of cumulative culture. It therefore stands to reason that at some point this becomes possible during human development. However, following the above logic, it is unlikely that we would find evidence of improvements with transmission in extremely young children whose ability to use social information is still developing (e.g. see the age-related improvements in performance documented by Atkinson et al., 2020, for even a trivially simple binomial discrimination task), at least within the context of the kinds of tasks typically presented to adult humans. Nevertheless, it is a reasonable simplifying assumption to consider that children’s capabilities typically increase with age, as a consequence of brain maturation and experience. Therefore, we would expect that the range of contexts within which cumulative culture might be manifested should become broader with increasing age. Based on what has been established to date in attempts to demonstrate cumulative culture in children, there is little that would appear to contradict this view, and some tentative evidence in support of it.

Three studies have attempted to test for a ratchet effect in children in ways that could potentially satisfy Mesoudi and Thornton’s (2018) core criteria. Reindl and Tennie (2018) adapted Caldwell and Millen’s (2008a) tower-building task (previously used to demonstrate cumulative culture in adults) for use with 4-5-year-old children. The participants’ goal was to

build as tall a tower as possible from sticks and plasticine. In transmission chains children were shown the tower built by the previous child or a replica of this. In contrast to Caldwell and Millen's (2008a) findings, there was no evidence of a ratchet effect, as tower height did not increase over generations within the chains. Nonetheless, children of this age appear to successively improve on others' solutions under different circumstances. Tennie et al. (2014) presented a task in which children were required to transport rice from one location to another, using any of a set of tools provided. Chains of 4-year-olds, who were able to observe the previous child's attempt before embarking on their own, were not able to improve on the efficiency of solutions that the children in the first generation (i.e. those with no exposure to social information) spontaneously adopted (Tennie et al., 2014). However, when the experimenter acted as the first member of the chain, and introduced an unnecessarily inefficient method, the children in some chains adopted more effective alternatives which were subsequently faithfully transmitted, leading to children in later generations transporting more rice. Similarly, Flynn (2008) found that chains of 2- and 3-year-old children improved the efficiency of tool-use behaviours, eliminating redundant elements that had been incorporated into the original demonstration by the experimenter. Both Tennie et al. (2014) and Flynn (2008) thus show only that children can improve upon abnormally suboptimal methods, as opposed to accumulating increasingly effective solutions – a process termed the “subtractive ratchet effect” (Tennie et al., 2014). Although this behaviour differs from that considered ratcheting by most researchers, these findings suggest that under certain circumstances ratcheting may be possible even in very young children. Creativity and/or insightful innovation, or technical reasoning skills (De Oliveira et al., 2019; Osiurak et al., 2020), may have been required for Reindl and Tennie's (2018) tower-building task. However, in other tasks improvements might readily arise as a consequence of copying error and/or random exploration. In such cases young children may derive benefits from the accumulation of task experience over multiple generations, relative to a baseline of no exposure to social information.

Other studies have tested children in groups, without generational turnover (Dean et al., 2012; McGuigan et al., 2017), claiming that beneficial innovations spread as a consequence of cumulative culture. However, in these studies it is possible that solutions became more effective due to children's increasing (direct) experience in the test situation: the benefits might not extend to situations involving generational turnover. This caveat aside, the findings are consistent with the idea that ratcheting may be possible from 3 years onwards, given a task for which the necessary innovations are relatively intuitive. The existing literature therefore suggests that the nature of the task may determine whether social transmission can lead to improvements in performance over multiple generations.

Considering the simplifying assumption that children's capabilities tend to increase with age, it follows that different tasks may have their own age threshold, at which evidence for cumulative culture can be identified.

To date, no study has directly compared age groups in relation to the cultural accumulation of task solutions. This omission is likely to be a consequence of pragmatic considerations, as running such a study using transmission chain or micro-society designs would be extremely challenging from a logistic perspective (Caldwell et al., 2020). Even when not comparing different populations, these methods require large numbers of participants. Each chain or micro-society (consisting of at least three individuals, but usually between five and ten) effectively represents a single sample unit. This therefore necessitates multiple chain/micro-society replicates in order to achieve statistical reliability, and to smooth out the effects of outlier performance by individual participants, which can disrupt an entire chain. Considering the challenges associated with recruitment of a narrowly-defined population (e.g. specific age bands), and the fact that these age bandings would likely need to be even more narrowly defined in order to generate a relatively clean comparison between groupings, it is little wonder few researchers have risen to the challenge.

In the current study we adopt an alternative approach to evaluating the potential for cumulative culture within specified populations, following logic outlined by Caldwell et al. (2020). Rather than relying on sample units consisting of multiple individuals (such as chains, replacement micro-societies, or closed group designs), and considering the effect of generation, or time, on the resulting performance, we test the potential for cumulative culture at the individual level, and consider the effect of exposure to task solutions of varying levels of success. The resulting pattern of performance can be used to infer whether, over multiple transmission episodes, this performance becomes more successful (if, for example – in the case of the most unambiguous evidence – participants were consistently outperforming the demonstration at all levels of task success).

Such an approach would not be possible using the kinds of tasks that researchers have, to date, used to study cumulative culture under laboratory conditions. In most of the experimental literature (and all developmental studies cited above), the method involves one single – usually quite complex – physical puzzle or problem. In these tasks, solutions can vary along a continuum of success, making it possible to track whether performance improves over generations of social transmission. However, using such an approach it is not possible to remove the influence of carry-over effects between multiple attempts and resulting from multiple demonstration exposures. Therefore, due to the inevitable order effects, it becomes impossible to determine how a particular individual would perform in the

absence of any other task experience, following exposure to demonstrations of varying levels of success.

Our approach uses a much simpler task (Caldwell et al., 2020) and relies on the fact that the information is strictly episodic, with demonstrations providing information that is specific to the subsequent test trial for that particular problem. Across different problems however, solutions are completely independent from one another. In addition, the nature of our task makes it possible to have precise control over the relative success of the demonstration. This allows us to present a single participant with multiple trials of the same task, without contamination of later trials as a consequence of prior exposure, and with the demonstrations potentially varying along the complete range of possible score outcomes. Thus, from testing at the individual level, we can infer the theoretical outcome of a linear transmission chain of similar performances, in a manner that would not be possible using established methods. For further details about this approach, see Caldwell et al. (2020).

We tested children using a stimulus selection task, a searching game involving an array of stimuli of varying reward value (in this case, simply rewarded or unrewarded), with the objective of maximising reward score. At baseline (i.e. without any social information) it is a guessing game, as stimulus features do not reveal their reward status. However, social information can be provided about reward value. This may require memorising the locations searched by a demonstrator, and whether or not these were rewarded. Alternatively, a demonstrator's activity may leave searched locations marked in some way, with this cue potentially also revealing information about reward value. Whilst the reward value of all stimuli is held constant between demonstration and the participant's selections for a given array (such that information acquired about the reward value of a stimulus during the demonstration can be assumed to hold true for the participant's attempt), each array has its own reward distribution. Therefore, as noted previously, it is possible to expose individual participants to multiple demonstrations of varying success (using different arrays) in order to infer the likely outcome of social transmission through a chain or populations of learners who behave like that particular individual. In addition to looking at children's reward score on the array as a whole, the stimulus selection task enables one to look at how a learner responded to social information regarding the individual rewarded and unrewarded locations selected within the array. This is informative because, if the social information on an array is to be fully utilised and offer the greatest performance benefit, a learner should respond very differently to these two types of information. Looking at this aspect of information use is a further advantage of the stimulus selection task compared to alternative methods in which it can be difficult to assess whether changes in behaviour occur primarily as a consequence of

imperfect replication, or active deviation from apparently ineffective elements of a demonstration.

The abstract nature of the stimulus selection task also allows us to manipulate task variables in precise and systematic ways. This is particularly valuable given the likely relationship between task context and the cognitive and behavioural traits required to generate a ratchet effect. In this study we manipulated the longevity of the cues provided by a social demonstration, creating very different demands in terms of memory load and environmental affordance. Cues either remained present in the environment or were present for a limited period – referred to as Enduring Cues or Transient Information conditions, respectively. The distinction between transient and enduring cues corresponds to differences in the information available as a consequence of others' activity in real world social information use. A conspecific's activity might, for example, expose a food source that was previously concealed (e.g. opening of milk bottles by birds, Fisher & Hinde, 1949), or generate partially processed food items that render the contents more accessible to an inexperienced individual (e.g. pine cone foraging in rats, Terkel, 1996). Potential foraging locations can also become marked as having been visited simply as a result of visitors leaving perceptible cues at exploited sites (e.g. honeybees and bumblebees leaving scent traces on flowers, Leadbeater & Chittka, 2007). In these examples, the information endures within the environment which removes both the need to witness the activity, and the potential cognitive burden of storing the information in either short-term memory or working memory, as would be required for transient information. Short-term memory has been defined as the storage of a limited amount of information (Cowan, 2008; Diamond, 2013), and working memory refers to both the storage and processing/manipulation of limited information (Best & Miller, 2010; Cowan, 2008; Diamond, 2013; Garon et al., 2008).

Contrast these examples with typical experimental paradigms used to study social learning (i.e. learning influenced by the observation of, interaction with or behaviour/actions of another animal, Heyes, 1994, 2012) in nonhuman primates and human children (e.g. see Whiten et al., 2009, for a review). These studies typically involve a single apparatus, operated by only one individual at a time, and the behaviour of interest is usually a specific method or action which can be performed on the apparatus. The social information available from others' interaction with the apparatus is therefore available only relatively instantaneously and leaves no lasting trace. Furthermore, such information requires storage since simultaneous activity is not possible.

The behavioural and cognitive requirements of social information use thus depend critically on contextual details such as those outlined above. Our aim was therefore to investigate the potential for ratcheting in children of a range of ages, under conditions of

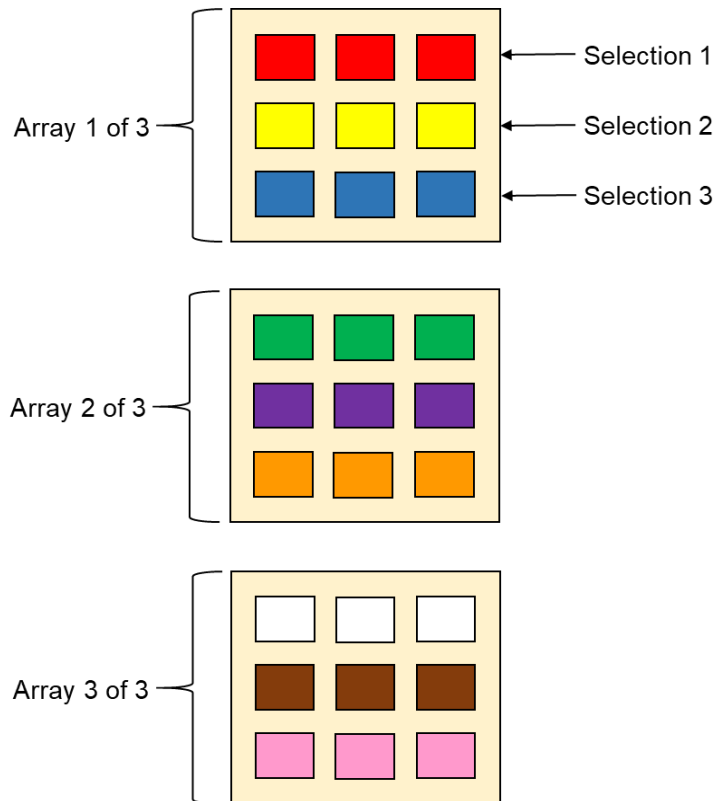
varying longevity of social information cues. Our expectation was that in certain populations (i.e. particular age groups, in this instance), a ratchet effect might be possible under conditions involving enduring environmental cues, whilst being precluded when cues were transient.

We presented the searching task to children by framing it as a challenge to find hidden treasure, a context we believed would be understandable and motivating for them. The reward (“treasure”) was always hidden in one of three locations (of the same colour, Figure 2.1A), and a total of three rewards could be searched for on each of three different, nine-chest arrays (Figure 2.1B). This meant that the total score on each array could vary between 0 (no rewards found) and 3 (all rewards found). We were thus able to create demonstration trials of varying success, and study how well children made use of social information about locations of rewards (arising from correct choices) and locations to be avoided (arising from incorrect choices). As well as predicting effects of age and cue longevity on overall task performance, we intended to use children’s scores on the different array types to infer the likely outcome of a series of social transmission events and assess the potential for ratcheting (in line with the logic in Caldwell et al., 2020). Children were therefore exposed to demonstrations reflecting three different success levels. In order to be classified as displaying the potential for ratcheting on this task, an individual (or in this case, age group) would need to perform significantly above chance level overall, showing some ability to benefit from the social information. However, the success of the individual or group in question would also need to be related to the success of the demonstration, such that higher scores followed higher scoring demonstrations. If demonstration success has no bearing on the child’s score then clearly no ratcheting would be possible, since social learners derive no additional value from exposure to beneficial modifications, resulting in a performance plateau. Indeed, assuming linear transmission, demonstrating a potential for ratcheting requires participant performance to be above the level of the demonstration for a minimum of two successive simulated transmission events. This criterion ensures that the evidence is consistent with not just benefits from exposure to social information, but also the accumulation of benefits over successive transmission episodes. In the current study therefore, our criteria for potential for ratcheting were outperformance of both a chance-level (1-rewarded) demonstration, equivalent to the typical outcome of naïve exploration, *and* an above-chance demonstration (2-rewarded).

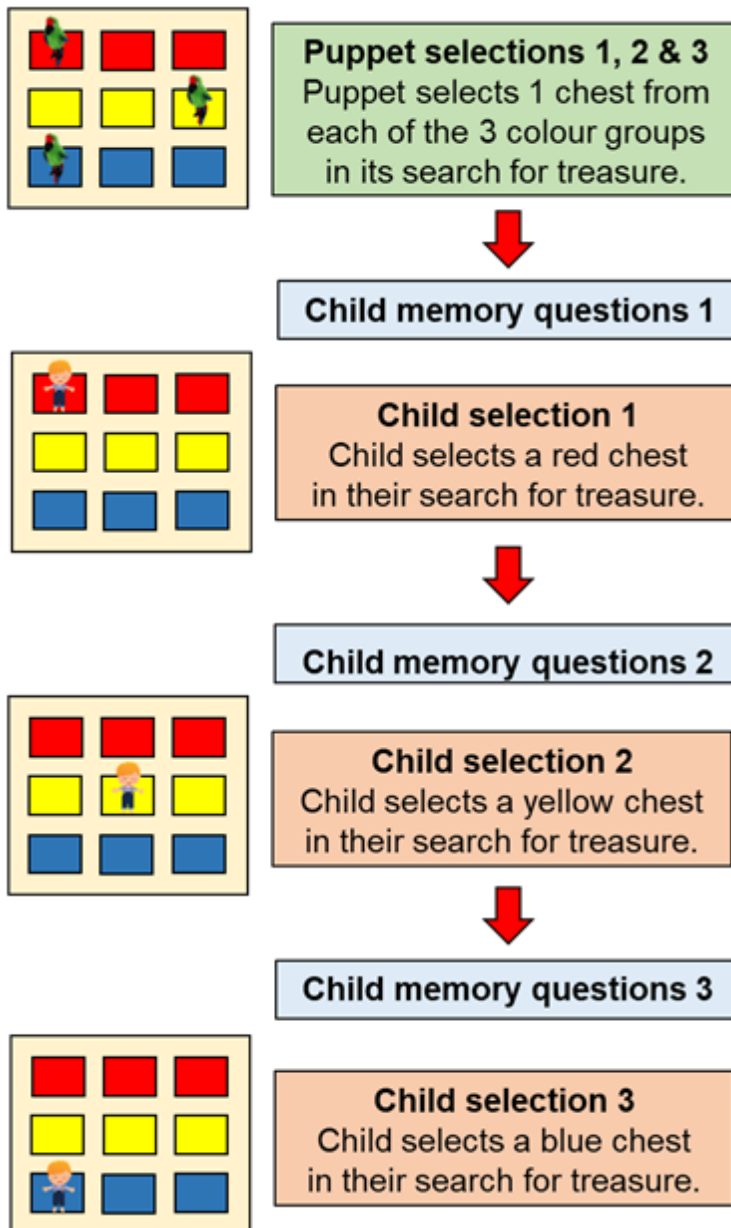
Figure 2.1

Diagram of Testing Procedures: Arrangement of Chests in Arrays 1, 2 & 3 (Panel A), and Summary of the Testing Procedure for Each Array, Using Array 1 as an Example (Panel B)

A.



B.



Note. Selections 1, 2 & 3 (labelled in Panel A) occur for each array.

We predicted that children would perform better in the Enduring Cues condition, compared with the Transient Information condition. As previously stated, using transient information is expected to place greater demands on memory. In order to achieve the best scores in our task we suspect that working memory may be utilised because both storage and manipulation of information are likely to be required (Best & Miller, 2010; Cowan, 2008; Diamond, 2013; Garon et al., 2008). In order to make use of information from unsuccessful

demonstration attempts (i.e. the revealing of unrewarded locations) one would presumably need to hold a memory of the unrewarded location in mind whilst deciding which of the remaining available locations to search, a similar process to that required in complex working memory tasks (Garon et al., 2008). We thus predicted that children would perform better with increasing age (in line with developments in working memory, Best & Miller, 2010; Diamond, 2013; Garon et al., 2008), across both conditions. Additionally, we used the criteria described above to estimate the extent to which age groups might contribute to the accumulation and retention of beneficial modifications through social transmission, under both conditions.

Materials and Methods

Participants

We collected data from 163 children aged 3 to 7 across the two conditions. The one 7-year-old was placed into the same category as 6-year-olds for descriptive statistics and t-tests but all further analysis used age in days. 83 children were recruited from Glasgow Science Centre and a further 80 from a primary school in Bradford, UK (further details in the Appendix). Nine children were excluded for the following reasons: missing date of birth from the consent form (age 3, female), failure to fully comply with task instructions (n=4; aged 3, 5 and two aged 4; male), inability to understand the task due to language difficulty (age 3, male) and experimenter error (n=3, aged 6 and two aged 4, female). The final sample consisted of 154 children aged 3-7 (M= 59 months, range= 38 – 84, SD= 13, 88 female). There were 76 and 78 children in the Enduring Cues and Transient Information conditions, respectively.

Ethics statement: This research was approved by the University of Stirling, General University Ethics Panel (references: GUEP40 & GUEP289). Written, informed consent was obtained from the parent or guardian of all children prior to their participation. Children were asked if they would like to participate, were continuously monitored for assent and were rewarded with a sticker regardless of task completion.

Apparatus

A large parrot hand puppet (The Puppet Company©) was used as a demonstrator, performing all social selections. The choice stimuli were wooden treasure-chests (82mm, 52mm, 46mm), containing either 30 (30mm x 30mm) squares of felt treasure (coloured to match the chests) as reward stimuli, or scrunched-up newspaper as a cue to the absence of a reward. Three chests of each of the following colours: red, yellow, blue, green, purple, orange, white, brown, and pink were presented in groups of three, nine-chest arrays (Figure 2.1A). Each array was placed in turn onto a laminated treasure map measuring 426mm x

482mm (Figure S.1). Two boards, each containing nine coloured squares (corresponding to the nine different colours of treasure to be found), were used to keep score (Figure S.2). Each time treasure was found by the puppet/child it was stuck to the corresponding colour on their score board.

Procedure

In the school, testing took place in a quiet room adjacent to the classroom. At Glasgow Science Centre, testing was carried out in a public space, separated from the main museum space. In the museum only, less confident children were accompanied in the testing area by a parent or guardian, who was instructed not to provide the child with any task-relevant assistance. In both locations the experimental task was carried out on a table-top with the child seated next to the experimenter and opposite the puppet (operated by a research assistant), Figure S.2. A verbal script was used by the experimenter (see Appendix “Experimenter Verbal Scripts”) and total testing duration was 15-20 minutes per child.

Children were asked if they would like to play a game in which the goal was to try to find more treasure than the puppet. A series of nine selections were made over three stimulus arrays (three selections per array, Figure 2.1B) and each selection was made from a choice of three chests of the same colour, which could be either rewarded (contain treasure) or unrewarded (not contain treasure). An array could therefore be viewed as a three-choice search task consisting of three puppet selections, and three child selections. In the Transient Information condition only, each child selection was preceded by two memory questions to assess memory of the puppet’s selections – note that the puppet performed all three of its selections before the child began their turn. The child’s turn then consisted of the following steps: memory questions for the puppet’s first selection, child’s first selection, memory questions for the puppet’s second selection, child’s second selection etc.

Each child was assigned to either a Transient Information or an Enduring Cues condition. In the Transient Information condition treasure chests were fully closed following the puppet’s selection, therefore information on the contents was only available for the duration of the choice itself. In comparison, chests in the Enduring Cues condition were left partially open so that the contents remained on view during the puppet’s remaining selections on that array, and throughout the child’s selections. The condition to which children were assigned was determined according to a combination of factors, including the maintenance of balanced numbers across conditions within age groups.

Children were first asked to look inside the three chests of the first colour group in the current array (e.g. red in Array 1 – Figures 2.1A and S.1) before placing them into a sack. They were then asked to do the same with the chests of the second and then third colour

groups. This was to ensure children understood that only one chest of each colour contained the treasure, and that the remaining two contained newspaper. The nine chests (of the three different colour groups in the current array) were then mixed up inside the sack so that the child could not know which chests contained the treasure. Next, the child was asked to remove the three chests of the first colour group (e.g. three red chests for Array 1) and place these on large Xs which were marked in a line at the top (left, centre and right) of a treasure map (Xs for the remaining two colour groups were in lines across the centre and bottom of the map) ready for the puppet to make its first selection, which it then did. Following this selection, chests of the second colour group were placed onto the map and so on until the puppet had made all three choices.

Puppet Selections. Each array consisted of three consecutive puppet selections which were made before the child made any selections. The assistant controlling the puppet knew which chests were rewarded due to discreet pencil marks drawn on the back of the chests. This enabled them to select a rewarded or unrewarded chest as predetermined. Prior to a puppet's selection, children were told that "Polly the Parrot" was going to look for the treasure (of one particular colour e.g. red for selection 1, Array 1) and the puppet then selected one chest by touching it with its beak; the child was prompted to open this chest. If the puppet was successful, the child was asked to take a piece of treasure and give it to the puppet, leaving the chest full of the remaining pieces of treasure. The puppet then placed its piece of treasure onto the relevant coloured square on its score board. If the puppet was unsuccessful nothing was removed from the chest. In the Transient Information condition chests were closed immediately following the puppet's selection, and in the Enduring Cues condition they were left partially open so that the rewarded/unrewarded contents were clearly visible.

Memory Questions. On each array, the three consecutive puppet selections were followed by three consecutive child selections. Each child selection was preceded by two memory questions (Figure 2.1B) – "Where did Polly the Parrot look for the [red] treasure?" and – "Did Polly find the [red] treasure?" Only their answer (given either verbally or by pointing) to the first question was used in the analysis because it was possible that children were using Polly's score board to determine whether she had found the treasure. These questions were asked in the Transient Information condition only and all children responded (see Appendix for the results).

Child Selections. Immediately following each set of memory questions children were asked to search for treasure in chests of the corresponding colour group. The procedure was as follows for each array: child memory questions 1, child selection 1, child memory questions 2, child selection 2, child memory questions 3 and child selection 3 (Figure 2.1B). Child selection involved a child choosing and opening one of the three chests in a colour group, as the puppet had done. If the child was successful, they were prompted to take a piece of treasure and place it onto their score board before the chest was closed. If they were unsuccessful nothing was removed from the chest and it was closed.

The total number of rewarded social selections in each of the three arrays differed. The puppet found one rewarded chest (in total across the three colour groups in an array) on one array, two rewarded chests on another, and three rewarded on a further array and thus scored 1/3, 2/3 and 3/3 – a total score of 6/9 which corresponded to six pieces of treasure found. The order of these three scores across the three arrays was counterbalanced: each of the six possible orders was assigned to participants within each age category in a randomised order. On the arrays in which 1/3 or 2/3 of the colour groups were rewarded we counterbalanced the position of the rewarded colour group(s) (1st, 2nd or 3rd) and assigned all possible combinations to participants within each age category in a randomised order.

The following responses were live coded: the puppet's selections (necessary because, although selection of a rewarded/unrewarded chest was predetermined, the exact unrewarded chest chosen was not), the child's responses to the two memory questions (although only answers to the first were used in the analysis, see Appendix), and the child's selections. Following a rewarded information trial, the optimal response was to repeat the puppet's selection and hence also find the treasure. Following an unrewarded information trial, the optimal response was not to repeat the puppet's response and to select one of the two alternative chests of the same colour, which resulted in a 50% chance of finding the treasure. If children always repeated the puppet's responses following rewarded trials, but shifted to different chests following unrewarded trials, we would expect an average child success score of 7.5. This would reflect six points for selecting the same chests as the puppet for the rewarded trials, plus an average score of 1.5 out of 3 for the unrewarded trials (50% correct).

Results

We were interested in the extent to which children used the social information provided by the puppet in our Transient Information and Enduring Cues conditions. We therefore measured information use with a number of dependent variables (discussed below). We were also interested in children's memory for the locations selected by the

puppet in the Transient Information condition (see Appendix). p -values $< .05$ were taken as statistically significant across all analyses. All generalised linear mixed effects models (GLMM) and generalised linear models (GLM) were carried out with either the logit link (binomial data, family = binomial) or log link (count data, family = poisson) and the lme4 package (lmer and glm functions for GLMMs and GLMs, respectively) using R (R Core Team, 2018). Our default choice for the random effects structure for each model included by-participant random slopes for variables which varied within participant, following Barr et al. (2013), to keep random effects structures “maximal” where possible. Where the “maximal” model resulted in non-convergent or singular fit models, random slopes were removed followed by random intercepts where necessary until a convergent, non-singular model was obtained. We list the final structure for each model, including all random intercepts and slopes, under the relevant sections below and in the Appendix (Table S.3). The following independent variables were used across our “Optimal Response Count” and “Repeating” analyses (below):

Between subjects – age (thousands of days, centred), condition (sum coded: Transient Information -1, Enduring Cues 1).

Within subjects – information type (sum coded, unrewarded set to -1, rewarded to 1).

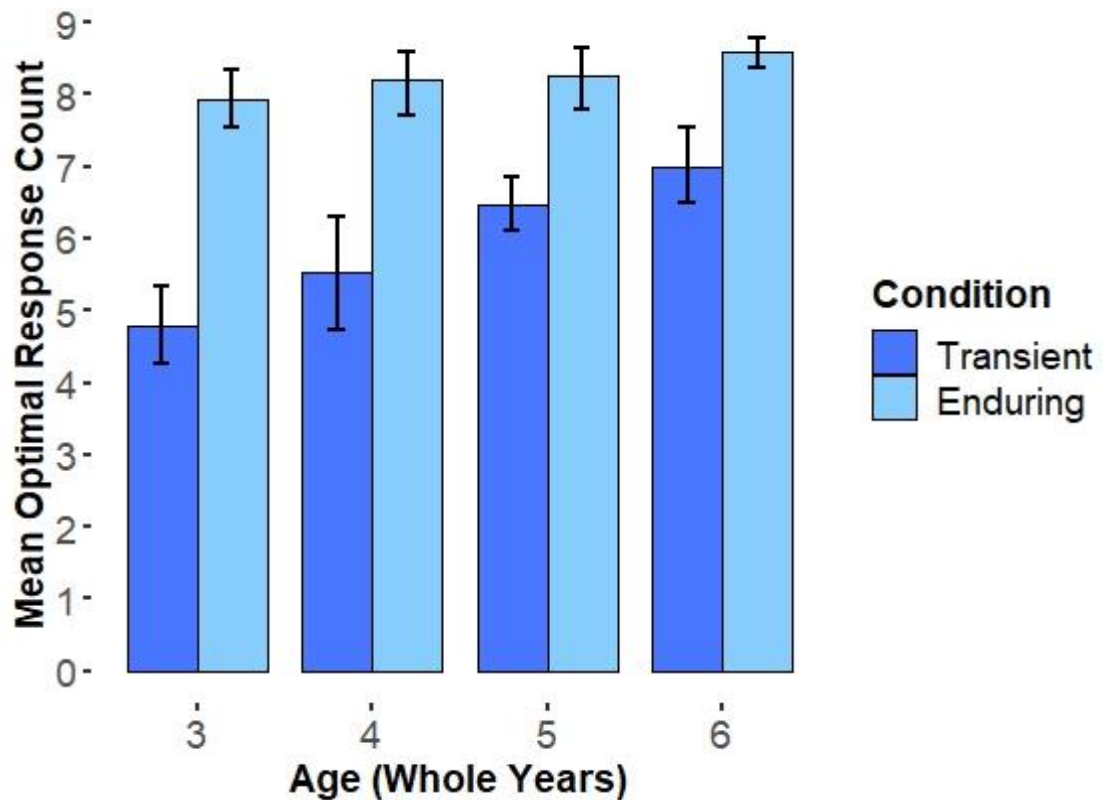
Optimal Response Count

In this analysis, children were awarded points for repeating rewarded selections, and not repeating unrewarded selections (see Appendix for further justification). As there were nine puppet selections, the maximum possible optimal response count was nine (Figure 2.2: mean counts by age and condition). Note that this count differed from a child’s success score on the game because, rather than simply being a sum of the number of pieces of treasure found, this accounted for the fact that it was possible to correctly not copy an unrewarded puppet selection and yet fail to find the treasure. However, despite giving a broad overview of information use, the optimal response count was a rather crude measure because it did not distinguish between what children did following rewarded and unrewarded social information trials, or following 1/3, 2/3 and 3/3 rewarded arrays.

We performed a generalised linear model with optimal response count as the dependent variable and age, condition, and the interaction of age and condition as fixed effects; there were no random effects. There were main effects of age ($p = .008$) and condition ($p < .001$): optimal response count was higher in older children and in the Enduring Cues condition (Figure 2.2). The interaction between age and condition was approaching significance ($p = .063$).

Figure 2.2

Mean Optimal Response Count/9 by Condition and Age (Whole Years)



Note. Error bars are 95% confidence intervals.

Despite the non-significant interaction effect between age and condition in our previous model, we further examined the effect of age on optimal response count in our two conditions. These analyses were considered worthwhile due to the relevance to our hypotheses of different age effects across the two task conditions. However, these models should be regarded as purely exploratory, and interpreted with caution accordingly. We again used a dependent variable of optimal response count, but condition was removed from the fixed effects and we split the data up by condition prior to running the models. These models therefore had age only as a fixed effect but were otherwise identical to the above. In the Enduring Cues condition, there was no main effect of age on optimal response count ($p = .539$). Contrastingly, there was a highly significant main effect of age in the Transient Information condition ($p = .003$): higher optimal response count in older children (Figure 2.2).

Repeating

Repeating the puppet's rewarded selections, and not repeating the puppet's unrewarded selections, are both correct strategies but pose different demands. While the former requires one to remember a rewarded selection and repeat it, the latter requires one to remember an unrewarded selection and avoid repeating it. If the social information was being used effectively, we would expect high and low levels of repeating following rewarded and unrewarded puppet selections, respectively. Repeating was binary coded, "1" for a repeat and "0" for no repeat, for all nine responses made by each child (see Table 2.1 for mean proportion of responses repeated by information type, age and condition and the Appendix for further details on this measure).

Table 2.1

Mean Proportion of Puppet Selections Repeated and Standard Deviation for all Ages Combined and Ages 3-6 (Whole Years) in the Transient Information and Enduring Cues Conditions, Rewarded and Unrewarded Information Types

<i>Age (whole years)</i>	<i>Information type</i>	<i>Condition</i>	<i>Mean proportion repeats</i>	<i>SD</i>
<i>All</i>	<i>Unrewarded</i>	<i>Transient</i>	<i>0.25</i>	<i>0.44</i>
<i>All</i>	<i>Unrewarded</i>	<i>Enduring</i>	<i>0.03</i>	<i>0.17</i>
<i>All</i>	<i>Rewarded</i>	<i>Transient</i>	<i>0.62</i>	<i>0.49</i>
<i>All</i>	<i>Rewarded</i>	<i>Enduring</i>	<i>0.89</i>	<i>0.31</i>
<i>3</i>	<i>Unrewarded</i>	<i>Transient</i>	<i>0.28</i>	<i>0.45</i>
<i>4</i>	<i>Unrewarded</i>	<i>Transient</i>	<i>0.28</i>	<i>0.45</i>
<i>5</i>	<i>Unrewarded</i>	<i>Transient</i>	<i>0.23</i>	<i>0.42</i>
<i>6</i>	<i>Unrewarded</i>	<i>Transient</i>	<i>0.22</i>	<i>0.42</i>
<i>3</i>	<i>Unrewarded</i>	<i>Enduring</i>	<i>0.02</i>	<i>0.14</i>
<i>4</i>	<i>Unrewarded</i>	<i>Enduring</i>	<i>0.03</i>	<i>0.18</i>
<i>5</i>	<i>Unrewarded</i>	<i>Enduring</i>	<i>0.05</i>	<i>0.22</i>
<i>6</i>	<i>Unrewarded</i>	<i>Enduring</i>	<i>0.02</i>	<i>0.13</i>
<i>3</i>	<i>Rewarded</i>	<i>Transient</i>	<i>0.44</i>	<i>0.50</i>
<i>4</i>	<i>Rewarded</i>	<i>Transient</i>	<i>0.56</i>	<i>0.50</i>
<i>5</i>	<i>Rewarded</i>	<i>Transient</i>	<i>0.69</i>	<i>0.46</i>
<i>6</i>	<i>Rewarded</i>	<i>Transient</i>	<i>0.78</i>	<i>0.42</i>
<i>3</i>	<i>Rewarded</i>	<i>Enduring</i>	<i>0.83</i>	<i>0.37</i>
<i>4</i>	<i>Rewarded</i>	<i>Enduring</i>	<i>0.88</i>	<i>0.32</i>
<i>5</i>	<i>Rewarded</i>	<i>Enduring</i>	<i>0.90</i>	<i>0.30</i>
<i>6</i>	<i>Rewarded</i>	<i>Enduring</i>	<i>0.94</i>	<i>0.24</i>

Effects of Information Type, Age, and Condition. We performed a GLMM with repetition of responses as the (binary) dependent variable; and information type, age, condition, and their interaction, as fixed effects. We included a by-participant random slope

for information type and random intercepts for participant, reward position (reward located on the left, centre or right of the row of chests) and trial number (1-9). There were main effects of information type ($p < .001$) and age ($p = .027$): more repeating following rewarded trials and in older children. There was no main effect of condition ($p = .098$) but there was a significant interaction between information type and condition ($p < .001$), showing that information type affected the likelihood of a response being repeated more in the enduring condition than the transient. There was no significant interaction effect between information type and age ($p = .051$), although this fell just short of statistical significance. There was no interaction between condition and age ($p = .648$) or information type, condition and age ($p = .263$).

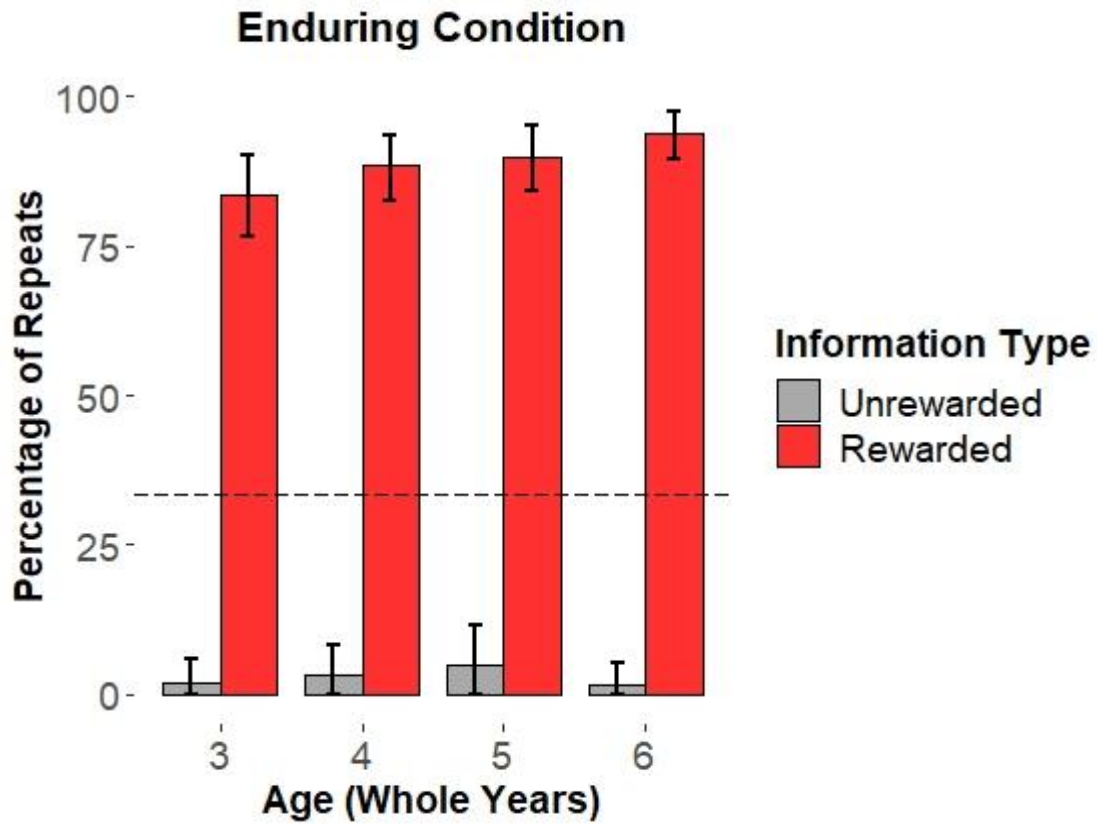
Effects of Information Type and Age in the Enduring Cues and Transient

Information Conditions. We ran two further models, with a view to understanding the differing effects of information type on repeating performance within each condition. These models had the dependent variable of repetition of responses (as above), but we removed condition from the fixed effects and instead split the data by condition. This left information type and age only as fixed effects. The model for the Enduring Cues condition had a by-participant random slope for information type and random intercepts for participant, reward position and trial number, as in the previous model. The model for the Transient Information condition had random intercepts for trial number and participant. We also added a random intercept for “remembered” (0 or 1 as per the child’s response to the first memory question) in order to account for any effect of ability to remember which chest the puppet selected on repeating. For the Enduring Cues we found a main effect of information type ($p < .001$): more repeating following rewarded trials than following unrewarded trials (Figure 2.3A). However, there was no evidence for an effect of age ($p = .400$), or an interaction between age and information type ($p = .768$). In the Transient Information condition (Figure 2.3B) there was also a main effect of information type ($p < .001$) and no main effect of age ($p = .470$). However, there was a significant interaction between age and information type ($p < .001$), which indicated that the effect of information type on repeating was stronger in the older children, compared with the younger. There was also a large variance (2.045) associated with the remembered variable.

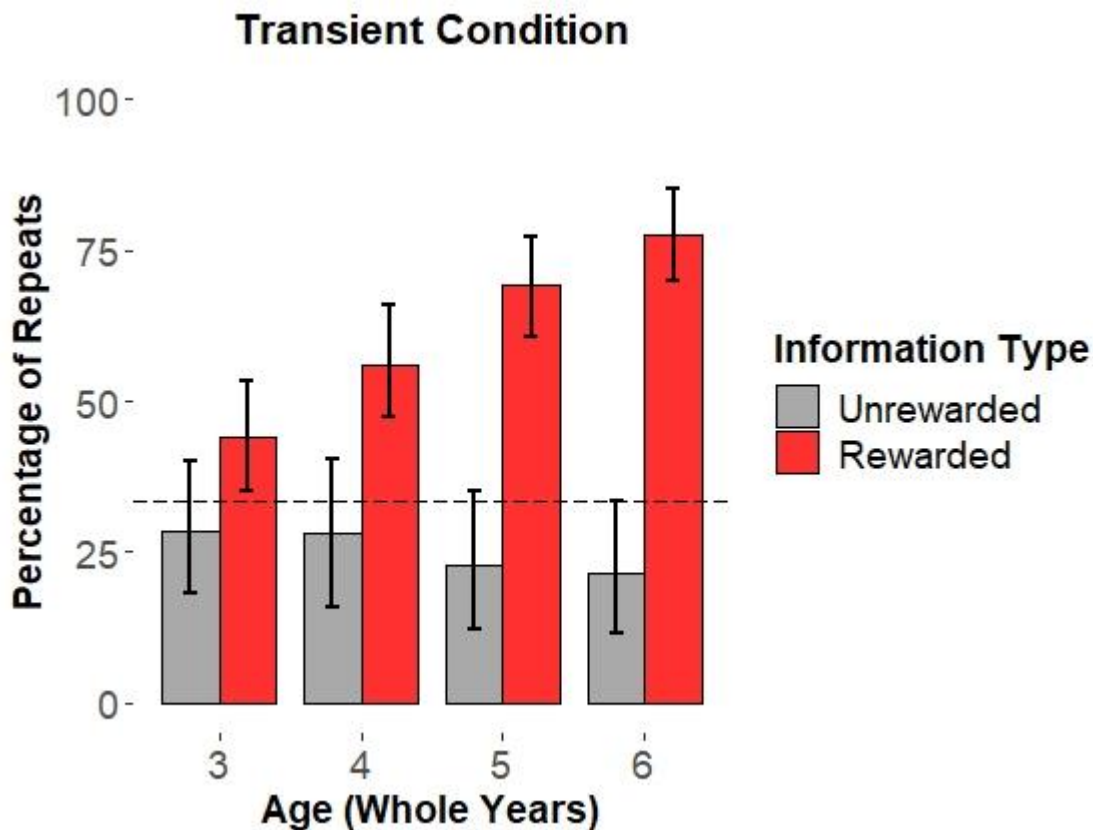
Figure 2.3

Mean Percentage of Responses in Which the Location Selected by the Puppet was Repeated for the Enduring Cues (Panel A) and Transient Information (Panel B) Conditions by Information Type and Age (Whole Years)

A.



B.



Note. Error bars are 95% confidence intervals. The dashed line shows chance performance: the proportion of repeats expected (for both unrewarded and rewarded problems) if children were selecting a chest at random on each turn and not using the social information.

Effects of Age and Condition by Information Type. We ran two further models, with a view to understanding the differing effects of age on repeating performance for rewarded and unrewarded information. We split the data by information type. These models therefore still had a dependent variable of repetition of responses but fixed effects of age and condition only. The model for the rewarded information had a random intercept for trial number and that for unrewarded had no random effects. For the rewarded selections we found main effects of age ($p < .001$) and condition ($p < .001$) – more repeats in older children, and in the Enduring Cues condition, but no interaction effects ($p = .431$). For unrewarded selections there was a main effect of condition ($p < .001$): fewer repeats in the Enduring Cues condition. However, there was no main effect of age ($p = .885$) or interaction of age and condition ($p = .377$). These models therefore showed that there was increased repeating with age for rewarded selections but not decreased repeating with age for

unrewarded selections despite both repeating after rewarded and not repeating after unrewarded being effective behaviours (we return to this point in our analysis of the second memory question, see Appendix). Overall repetition of responses (more repeats after rewarded selections and fewer repeats after unrewarded selections) was more accurate in the Enduring Cues condition (Figure 2.3A) in which the information was more freely assessable.

Potential for Ratcheting (PFR¹)

We grouped children according to chronological age, using one-year bandings, and classified the performance of each group, in each condition, according to the PFR continuum described below. For an age group to be defined as showing PFR in a linear transmission chain the Level 3 criteria needed to be fulfilled.

Our PFR classification continuum was as follows: Level 0: Chance-level performance across all trials, regardless of demonstration success (no benefit from social information); Level 1: Above chance performance overall, but scores no higher, on average, following higher success demonstrations (2/3 and 3/3), compared with lower (1/3) (social learning benefit unlinked to demonstration success); Level 2: Above chance performance overall, with scores higher, on average, following higher success demonstrations (2/3 and 3/3), compared with lower (1/3), without outperforming higher success demonstrations (2/3) (social learning benefit linked to demonstration success); Level 3: Outperform demonstrations by scoring higher following both chance-level (1/3) and higher success (2/3) demonstrations (potentially supporting ratcheting under linear transmission).

For Level 3 ratcheting, we included the first of these criteria because outperforming a chance-level (1/3) demonstration shows that the ratchet effect can get off the ground, with naïve exploration as a starting point. The second criterion verifies that subsequent transmission events can result in further performance improvements, by evaluating whether social learners also outperform a demonstration equivalent to expected optimal performance resulting from exposure to naïve exploration (2/3). This pattern of performance would support ratcheting even under linear (single cultural parent) transmission. However, under some circumstances, ratcheting may occur without outperformance of an individual model's higher success demonstrations. For example, in a situation in which there is more than one cultural model, performance may improve following the use of multiple pieces of valuable information from different models. Such non-linear transmission scenarios are captured by our Level 2 criteria. Under these Level 2 criteria, it had to be clear that children were performing better following high-success demonstrations (2/3 and 3/3), compared with lower-

¹ We hereafter use the abbreviation "PFR" to denote "Potential for Ratcheting".

success demonstrations (1/3), illustrating that success was linked to the quality of information available. Although children achieving this level were not able to outperform the higher-level 2/3 demonstration directly, the benefit they got from higher success demonstrations means that, at a population level, performance could potentially improve over generations if learners were exposed to multiple demonstrations.

In order to classify children of each age group according to our PFR continuum we gave children three success scores, each representing their aggregated performance on one of the different success level arrays (1/3, 2/3 and 3/3 rewarded). Children were given a point each time they found the treasure and zero points if they did not. They could therefore score up to one point on each colour group, a maximum of three points on each array and nine points over the entire task. We calculated the mean success scores for each age group and condition for the entire task (success score/9); the one-, two-, and three-rewarded arrays; and the mean of the two and three rewarded (Table 2.2 and Figure 2.4). We then performed one-tailed *t* tests and used these results to classify the performance of each group according to our PFR continuum (Figure 2.5). We employed a “strict” and a “less strict” method when assessing whether children fulfilled the requirements necessary to be categorised as a particular level. For the “strict” criteria, children from a particular age group needed a mean success score which was significantly higher than the benchmark outlined in a particular level criterion but for the “less strict” criteria a numerical difference, but not significance, was required. We report the less strict criteria in full here, due to the crossing (or otherwise) of the threshold being perhaps more relevant for this purpose than the degree of confidence in that conclusion. However, results according to the strict criteria are reported in full in the Appendix, and both are displayed on Figure 2.5.

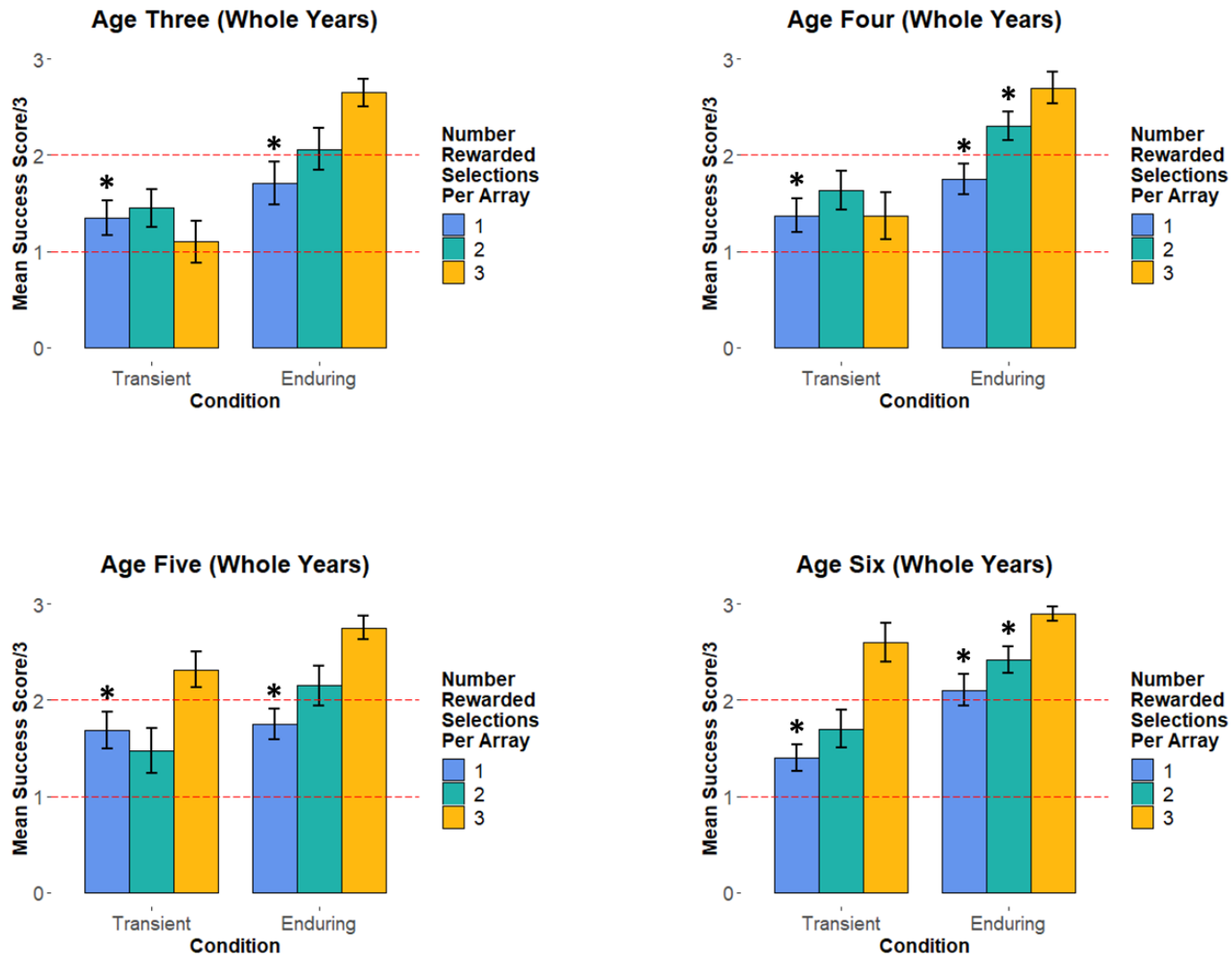
Table 2.2

*Mean and Standard Deviation for Success Score/9 and Success Score/3 for Ages 3-6
(Whole Years) in the Transient Information and Enduring Cues Conditions*

Age (whole years)	Condition	Success score/9		Success score/3							
				1 rewarded		2 rewarded		3 rewarded		2 and 3 rewarded	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
3	Enduring	6.41	1.33	1.71	0.92	2.06	0.90	2.65	0.61	2.35	0.52
4	Enduring	6.75	1.29	1.75	0.72	2.30	0.66	2.70	0.73	2.50	0.46
5	Enduring	6.65	1.35	1.75	0.72	2.15	0.93	2.75	0.55	2.45	0.54
6	Enduring	7.42	0.90	2.12	0.74	2.42	0.61	2.89	0.32	2.66	0.34
3	Transient	3.90	1.37	1.35	0.81	1.45	0.89	1.10	0.97	1.28	0.55
4	Transient	4.37	1.83	1.37	0.76	1.63	0.90	1.37	1.07	1.50	0.75
5	Transient	5.47	1.31	1.68	0.82	1.47	1.02	2.32	0.82	1.89	0.59
6	Transient	5.70	1.26	1.40	0.60	1.70	0.86	2.60	0.88	2.15	0.49

Figure 2.4

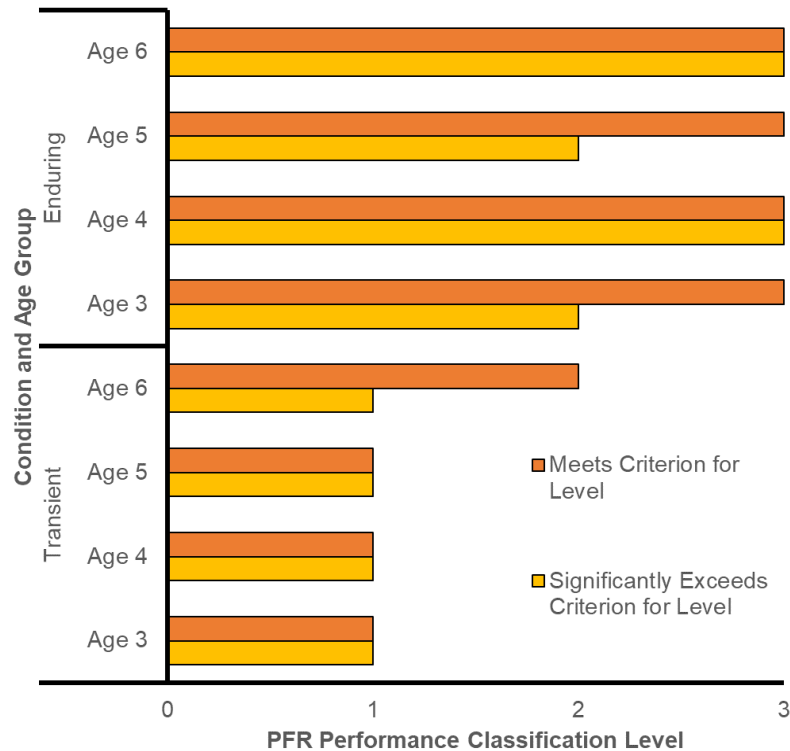
Mean Success Score/3 by Number of Rewarded Selections in an Array and Condition for Ages 3, 4, 5 and 6 (Whole Years)



Note. The dashed line at 1 depicts chance performance: the score expected if children were selecting a chest at random on each turn and not using the social information. The dashed line at 2 allows visualisation of whether children scored above 2 for the two-rewarded arrays. An asterisk indicates that the mean score is significantly above 1 (chance) or 2 for one- and two-rewarded arrays respectively. An asterisk above both bars denoting one- and two-rewarded arrays, within an age group and condition, illustrates Level 3 ratcheting. Error bars indicate \pm the standard error of the mean.

Figure 2.5

PFR Performance Classification Level by Condition and Age (Whole Years)



Note. Situations in which a mean success score was *significantly* higher (yellow bars) and *higher* but not *significantly* higher (orange bars) than the benchmark outlined in a particular level criterion are described using the terms “Significantly Exceeds Criterion for Level” and “Meets Criterion for Level” respectively.

According to the less strict criteria, in the Enduring Cues condition, children of all ages were categorised as Level 3 (Figure 2.5) because the mean success scores for the 1/3 demonstration were greater than chance-level (1) and the mean success scores for the 2/3 demonstration were greater (although not all *significantly* greater) than 2 (Table 2.2 and Figure 2.4). However, in the Transient Information condition, although the mean success scores were greater than chance-level for each age group, the mean success scores for the 2/3 demonstrations were not >2 for any age group (Table 2.2 and Figure 2.4). Furthermore, mean success scores for the 2/3 and 3/3 demonstration types combined (Table 2.2) were above 2 at age 6 only; therefore children aged 3, 4 and 5 were classified as Level 1 and children aged 6 as Level 2 in the Transient Information condition (Figure 2.5).

Discussion

We used a novel method of testing capacities for cumulative culture, which allowed us to evaluate performance at the individual level. This gave much greater sensitivity for identifying age effects in children and was motivated by our aim of testing our ideas about the constraints on cumulative culture. Specifically, we had reasoned that whether cumulative culture would be manifested within a given population depends on the demands of social information use in the specific learning context in question, as well as the characteristics of members of the population. It followed that, in certain populations, performances indicative of the potential for cumulative culture might be observed in some contexts but not others. In line with this reasoning, we had predicted that children's age would be positively related to their scores on our proxy measures of potential for ratcheting. We also expected that, in a version of the task with reduced cognitive demands, scores would be higher than they were in an otherwise similar task that imposed significant demands on memory. This was expected to result in patterns of performance indicative of potential for cumulative culture being more prevalent under reduced cognitive demands, and apparent from a younger age.

Our results broadly supported these predictions, in that scores were higher in our Enduring Cues condition relative to the Transient Information. Furthermore, in the latter it was only in the oldest age group (6 years) that we found any suggestion of the potential for a ratchet effect and our optimal response count increased with age in this condition only. When looking at information type separately we found that children's ability to utilise information about rewarded selections increased with age, but not for unrewarded, which did not significantly improve (we discuss how this may relate to the result of the first memory question in the Appendix). We return to evaluate the rationale and effectiveness of these metrics later in the discussion, and first consider the validity of the contrasting conditions, and the distinction between them.

It is important to note that in our Enduring Cues condition, which left visible evidence of a location's reward value, the "demonstration" element (i.e. observing the puppet's choices) is likely redundant, or influential only inasmuch as it set the information in context. In this condition the demonstration may have functioned to reinforce understanding of the goal of the task and the desirability of the contents of the boxes (coloured pieces of "treasure" versus crumpled newspaper). However, assuming this understanding was in place, the knowledge that the cues arose as a consequence of another's activity was not required in order for the children to make effective use of the information. Arguably therefore, the Enduring Cues condition did not involve "social" learning at all. It is likely that the children would have performed just as well, or possibly better, had the cues been generated by their

own individual exploration, rather than the social demonstration. This applies to both conditions.

Other work in our group (Atkinson et al., 2020; Renner et al., submitted) has investigated children's performance following exposure to information acquired from a social demonstration compared with individual exploration, in logically similar – although simplified – tasks. Both found that children's use of information was largely unaffected by source. However, in the current study, we were not concerned with the question of children's understanding of the source of the information, nor whether their responses might be dependent on the nature of that source. Here we simply wished to evaluate their ability to use such information in ways which could potentially lead to ratcheting of task performance based on cultural transmission, and the conditions under which this might be possible. However, as noted on page 35, the definitions of cumulative culture used by some researchers are more restrictive than ours, which adhered to Mesoudi and Thornton's (2018) core criteria. Such researchers may include additional requirements (e.g. Mesoudi & Thornton's, 2018, extended criteria) and thus not regard our task as suitable for detecting cumulative culture. This is consistent with our view that the cognitive demands associated with cumulative culture may differ depending on the precise context of the behaviour being transmitted – more restrictive definitions of cumulative culture likely involve situations with a different set of cognitive demands.

As outlined in our introduction, there are now examples of information ratcheting in migratory species previously unstudied in relation to cumulative culture – pigeons (Sasaki & Biro, 2017) and bighorn sheep and moose (Jesmer et al., 2018). In both, benefits to subsequent generations are transmitted by virtue of the grouping tendencies of the species in question, and those benefits likely accumulate as a consequence of direct feedback to the individuals based on relatively random deviations from the acquired behaviour. Therefore, there appears to be no necessity for complex cognitive abilities to either extract cloaked information or to innovate new solutions. We consider these examples to be somewhat comparable to our Enduring Cues condition in which children had full access to information which could be utilised to improve their success score. Our results allow us to infer what would occur in a social transmission scenario involving a series of children attempting the task once each in succession. Results from the Enduring Cues condition suggest that children would typically achieve maximum score (finding all rewards) following just a few generations of transmission (probably by the fourth or fifth child to attempt the task). In contrast, results from the Transient Information condition suggest that information about reward locations would not be reliably retained during transmission, preventing accumulation of benefits. If cognitive abilities not within a child's or species' repertoire are needed in order

to access certain information (e.g. memory ability in our Transient Information condition) then this places limits on the potential for ratcheting over generations. However, it follows that ratcheting may be perfectly possible when such constraints are absent. Our finding that other aspects of cognition (i.e. not specifically socio-cognitive) may place constraints on the potential for cumulative culture is consistent with emerging literature. For example, Osiurak and Reynaud (2020) have highlighted the importance of technical reasoning skills for cumulative culture of technology, and Tennie and colleagues (e.g. Tennie et al., 2009) have proposed the “zone of latent solutions” hypothesis which posits that limits on nonhumans’ individual learning capacities will place constraints on their cultural behaviours.

It has been proposed that enhanced working memory in the human lineage has enabled cumulative culture, however this opinion is based on evidence from the archaeological record rather than measured differences in cognition (Balter, 2010; Coolidge & Wynn, 2001; Wynn & Coolidge, 2011). Comparing human and nonhuman memory directly remains challenging due to the small amount of comparative research available for analysis. It is likely that some nonhuman mammals have a similar memory storage capacity to that of humans, but that humans can represent concepts within memory differently and have a better ability to deploy attention and resist interference (for a systematic analysis of evidence to date see Carruthers, 2013). Humans can also extend their own memory capacity through the development of external storage systems such as writing (which acts to enhance working memory through literacy learning, see Wynn & Coolidge, 2011), mnemonic devices (Jurowski et al., 2015) or inner speech which allows for the rehearsal of items held in memory (Carruthers, 2013; Cowan, 2008). Such strategies have been described as “cognitive offloading” and children’s tendency to use them increases with age (Armitage et al., 2020). It may therefore be the case that young children, who are yet to fully develop metacognitive storage strategies or the ability to devise cognitive offloading strategies (Bulley et al., 2020), have a greater reliance on their raw memory storage capacity when solving tasks such as that in this study, and may therefore perform similarly to some nonhumans. If nonhuman primates were to complete a similar task, assuming the relevance of the cues had been learned, we would expect them to show PFR in the Enduring Cues (as children did from age 3) due to the low memory load requirements. Comparatively, if performance on the Transient Cues relies on using strategies such as inner speech, or if humans do have greater memory capacities, it is unlikely that other species would show PFR in this condition.

In this study we used a simple stimulus selection task to minimize incidental task demands that might constrain performance, and to allow us to manipulate key variables of interest. This extreme simplification means that the task probably deviates in significant

ways from most real-world cases of social learning. For example, the “decision” as to whether the social information should be repeated is binary: either repeat the demonstrator’s choice or choose an alternative location. Furthermore, the payoffs of the demonstrator’s choices are completely transparent, and the structure of the reward landscape is known (i.e. one of the three options is rewarded), even if the location of the reward is not. Therefore, it is possible to work out the optimal response, subject only to the availability of the information either in the environment or from internal storage. Whilst it could be argued that these simplifications are unrealistic, and that our paradigm reveals little about social learning in the real world, we contend that the demands we eliminated (opaque payoff information, ill-defined solution space, and unknown reward landscape) are not peculiar to social learning. There is therefore no reason that social learning paradigms must encompass these associated challenges. Task simplification allowed us to establish that a ratchet effect was possible, in principle, even in very young children. However, the stimulus selection task also offers scope to investigate alternative questions related to social information use e.g. ritual or strategy use (Kapitány et al., 2018).

Based on children’s PFR performance in our Enduring Cues and Transient Information conditions it appears that superior cognitive capacities (such as working memory) were required to access (and therefore use) the information provided in the Transient Information condition compared to the more accessible information in the Enduring Cues. We conclude that the potential for a ratchet effect (whether considering children of a particular age, or members of a nonhuman species) depends on the type of information itself, as well as the cognitive capacities required to access that information.

Funding: This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme. Grant agreement No. 648841 RATCHETCOG ERC-2014-CoG.

Chapter 3: Children's Use of Social Information From Multiple Models: Cognitive Capacities Underlying Population Size Effects on Cumulative Culture

Note. This chapter has been submitted for publication in its current form:

Wilks, C.E.H., Atkinson, M., & Caldwell, C.A. (submitted). Children's Use of Social Information From Multiple Models: Cognitive Capacities Underlying Population Size Effects on Cumulative Culture. *Culture and Evolution*.

Abstract

Population size has been proposed to promote cumulative culture in humans. Experimental evidence from adult humans suggests that this may be due to the potential for combining beneficial information from multiple models. However, it is possible that such combinatory social learning requires cognitive capacities restricted to adult humans. In our task, children aged 5-10 were exposed to two models who consecutively searched a 3x3 array for rewards. Models revealed different correct and incorrect reward locations. This information could be used by the child to maximise their own score on the same task. We were interested in children's ability to select rewarded locations, and avoid unrewarded ones, revealed by both models. We also manipulated the spatial and temporal displacement of the information available. Results showed that the youngest children were unable to fully benefit from the additional information provided by the two models under spatial and/or temporal displacement. Such displacement likely applies in most real-world cases of cumulative culture therefore our result may offer insight into the constraints on cumulative culture in nonhumans.

Introduction

The human propensity to continuously improve our cultural products over generations of learners has been proposed to be the result of increases in population size, rather than changes in cognition (Collard et al., 2013; Henrich, 2004; Kline & Boyd, 2010; Kobayashi & Aoki, 2012; Powell et al., 2009; Shennan, 2001). It has been suggested that individual learners utilise the additional information from larger populations by selecting the best of a larger number of individuals from whom to learn (Henrich, 2004), or by combining information from multiple different individuals (Kempe & Mesoudi, 2014; Muthukrishna et al., 2014). However, using such mechanisms may depend on cognitive capacities unique to adult humans. If this is the case, then human cognition may still form part of the explanation as to why increased population size appears linked to vast cultural expansion in humans and not in nonhumans.

We focus on the ability to combine information from multiple models, a process likely to involve integration of information separated in time and/or space (Subiaul et al., 2015). For example, integration over time might be required for learning a foraging skill; this could involve combining information from observation of an individual extracting the contents of a shelled plant, with information obtained through a separate observation of the technique required to break into this item. To use the same example, integration over space (but not time) would be involved if the learner (with access to their own fruit or nut, allowing simultaneous activity) could observe two individuals concurrently, one of whom was demonstrating the opening technique, and the other the extraction. It is probably fair to assume that such integration is within the cognitive abilities of most adult humans. However, there is some reason to doubt the capabilities of human children in this regard. If integration of information under temporal and/or spatial displacement does pose a significant cognitive challenge, then we would expect children to become better at utilising information from multiple models with increasing age. Furthermore, information about how performance changes with age may provide insights into the specific cognitive abilities that may be involved, and by extension, whether such capacities are likely to be available to nonhumans. We thus sought to investigate this experimentally, across a broad developmental range, in conditions in which the information from multiple models was presented differently in time and/or space.

As humans, our ability to improve and build upon our prior achievements is unparalleled and ubiquitous – evidenced in our languages, complex technologies, societal structures and ability to exploit most environments on the planet. These are examples of what has been termed cumulative culture, a process whereby a directional pattern of change

results in “improvements” (Tennie et al., 2009) or increasingly “preferred” traits (Caldwell, 2018) as behaviours or cultural products are transmitted over generations of social learners (Mesoudi & Thornton, 2018). This notion of constant improvement has led to cumulative culture also being referred to as the *ratchet effect* (Tennie et al., 2009; Tomasello, 1990). We use these terms interchangeably.

Recent evidence suggests that, at least according to some definitions (Mesoudi & Thornton, 2018), cumulative culture may exist in nonhumans (Jesmer et al., 2018; Sasaki & Biro, 2017). However, it is undeniable that this has not constituted anything on the same scale as the phenomena observed across human societies. There are many factors proposed to have driven this cultural expansion in humans (Dean et al., 2014); the factor we focus on here is demography. The results of theoretical models (Henrich, 2004; Kobayashi & Aoki, 2012; Powell et al., 2009; Shennan, 2001), supported by some ethnological studies (Collard et al., 2013; Kline & Boyd, 2010), suggest that changes in population size, rather than cognition, could have accounted for periods of rapid cultural expansion (e.g. during the upper Palaeolithic 45000 years ago, Powell et al., 2009; Shennan, 2001) or loss (as in “The Tasmanian Case”, Henrich, 2004) in human groups. Although cited widely, these conclusions have also been fiercely debated (Collard et al., 2016; Henrich et al., 2016; Vaesen et al., 2016).

Theoretical models (Henrich, 2004; Powell et al., 2009; Shennan, 2001) have investigated conditions under which larger numbers of interacting individuals (an “effective” population) can support behaviours or products of increased cultural complexity/functionality. For example, Henrich’s (2004) model (extended by Kobayashi & Aoki, 2012) assumed that learners were able to selectively copy the individual with the best available variant of a cultural trait (success-biased copying), but that copying was inaccurate so these inaccuracies would, under certain conditions, ultimately lead to a loss of skill within the population. However, larger populations had an advantage; they were able to negate this loss because they contained more individuals with traits which were of higher than average fitness hence a learner was more likely to select one of these more successful individuals from whom to learn. This was found to counteract the negative effects of copying (low transmission fidelity), leading to cumulative culture in populations above a certain size threshold.

It is important to note that the above models assume the success-biased copying (see Wood et al., 2013, for a review of such biased social learning strategies) of randomly derived higher fitness traits, although Henrich (2004) states that some improved trait variants may derive from individual learning (also considered novel invention, Lewis & Laland, 2012). However, based on further modelling work, some authors have proposed that combining

information from multiple different models is likely to constitute the main method by which improved trait variants arise (Lewis & Laland, 2012) and that a cultural trait is more likely to become established if it is learnt from multiple models (Enquist et al., 2010). Relatedly, innovation through combination has been proposed to be an important mechanism underlying the evolution of technology (see Winters, 2020, for examination of the importance of this mechanism versus minimisation of information loss through the use of social learning mechanisms).

Increasingly therefore, theorists have argued that the production of improved variants (that is, trait variants which represent an improvement or upgrade on those previously found in a population) through novel invention may play a relatively insignificant role in advancing cumulative culture (Lewis & Laland, 2012), and that the popular supposition of “genius” inventors should perhaps be reconsidered (Muthukrishna & Henrich, 2016). Evidence consistent with the importance of recombination has also been identified in the form of phylogenetic analyses on the constituent elements of technologies; existing forms of devices, such as radios (O’Brien & Lyman, 2000) or bikes (Lake & Venti, 2009), have been combined to produce the latest versions (Muthukrishna & Henrich, 2016). Additionally, there is now experimental evidence (Kempe & Mesoudi, 2014; Muthukrishna et al., 2014) that adult learners with access to multiple models combine information across models to generate better solutions than those with access to one model.

As outlined above, we take an experimental approach to investigating the integration of information across multiple models in children. As such, we will now discuss the experimental evidence to date in detail, beginning with relevant work performed in adult populations. Muthukrishna et al. (2014) tasked participants in transmission chains to recreate a complex image using image editing software. The transmission chains were ten-generations long and were structured such that participants were exposed to information from either one or five models from the previous generation. Each participant provided a screenshot and written instructions to help those in the subsequent generation. There was evidence of cumulative culture in the five-model group only, and furthermore, participants appeared to preferentially utilise information from the top performing model, in addition to the next three top performers, to generate novel recombinations possessed by none of the cultural parents.

Kempe and Mesoudi (2014) ran transmission chains, with either one or three models in each of four generations. The task was to complete a 100-piece jigsaw puzzle and participants had full access to the attempts of the participant(s) from the previous generation. As in Muthukrishna et al. (2014), there was evidence of ratcheting in the group with multiple models only (comprising five separate chains), measured as an increase in the mean

number of puzzle pieces correctly connected to other pieces as generation increased. It appears that this ratcheting was possible because the presence of multiple models provided an increased amount of information which could be integrated. This was apparent from the number of unique puzzle pieces correctly connected by all three models within a single generation for any given chain, i.e. the total package of information about the puzzle which was passed to the next generation.

Derex et al. (2013) has also performed experimental work which supports the effect of model number on cumulative culture. The relationship between group size and cultural complexity was investigated through introducing a simple and a complex artefact building task into groups containing two, four, eight or 16 members (players in a computer game). Participants were constrained to selecting to learn from just one of these players from the previous generation and thus this study is not directly relevant to our discussion of ability to combine information from multiple models. However, making a judgement as to which of a number of demonstrations is more valuable, and therefore would be preferable to utilise, may involve cognitive challenges similar to those required when combining information from more than one demonstration.

The above experimental work thus appears to support the relationship between population size and cultural change as described by the evolutionary models discussed (Henrich, 2004; Powell et al., 2009; Shennan, 2001). Importantly, these studies have highlighted the value of additional research into the underlying mechanisms, particularly the role of recombination (Kempe & Mesoudi, 2014; Muthukrishna et al., 2014), perhaps in addition to success-biased copying (Derex et al., 2013), in creating improved trait variants.

However, irrespective of the mechanism underlying the effect of population size on cumulative culture, it is unlikely that more models would result in cumulative culture without additional cognitive abilities enabling individuals to exploit and integrate information from multiple models. Thus, population size alone is unlikely to be responsible for the observed differences in cumulative culture expression between humans and nonhumans (Dean et al., 2014) and cognitive abilities unique to human adults may also be necessary. Both mechanisms suggested to underlie the population-size effect – utilising information from the “best” individual (success-biased copying, Henrich, 2004; Wood et al., 2013) or combining information presented by multiple individuals (Kempe & Mesoudi, 2014) – presumably involve additional cognitive demands such as memory, spatial translation and evaluation of alternative information (we discuss these further below). If cognitive abilities unique to humans, or used differently in humans, are implicated in overcoming these challenges then this may contribute to the explanation as to why increases in population size appear linked to the expansion of cumulative culture in humans. Research into the possible role of cognition

has however been largely neglected by researchers studying population size and cumulative culture. Consequently, we considered this a fruitful avenue for research.

We now discuss two further adult experimental studies which suggest that aptly integrating information from multiple models may be cognitively demanding, such that benefits cannot be guaranteed under all circumstances, and may even be *reduced* by the increased cognitive burden. Caldwell and Millen (2010) and Fay et al. (2019) investigated the relationship between population size and cumulative culture using transmission chains in which the participants' goal was to build paper aeroplanes which could fly as far as possible. Both studies included conditions in which participants had access to information from different numbers of models (one, two or three in Caldwell & Millen, 2010, and one, two or four in Fay et al., 2019); the study by Fay et al. (2019) also included an individual learning condition (repeated attempts replacing generational turnover). Rather surprisingly, Caldwell and Millen (2010) found cumulative improvements in plane flight distance as generation number increased for the one and two model conditions, but not the three model. Similarly, in Fay et al. (2019) flight distance increased with generation number for the individual learning and one model conditions, but not the two and four model conditions. In Fay et al. (2019) the previous generations' planes (e.g. four in the four-model condition) were only available to view (individually) for a short period prior to a participant attempting to build their own plane. In Caldwell and Millen (2010), although some of the previously built planes were on display during a participant's building time (e.g. three in the three-model condition), the time available to view each individual plane in the three-model condition was extremely short, with each removed after only a brief exposure time. Participants therefore did not have continuous access to all of this potentially valuable information for the full building period. In these studies, it therefore seems probable that processing the different models' designs and keeping them in working memory in order to select the best design (or make use of the beneficial elements from previously observed designs) became more challenging as the number of models increased. That is, once the number of models increased beyond a certain point, cognitive demands associated with integrating the information constrained its usefulness. These studies therefore support the theory that using social information from multiple models may pose significant cognitive challenges, which we will now discuss further.

Utilising social information, whether from one or multiple models, may give rise to cognitive challenges due to the way in which the information is presented. Our previous work (Wilks et al., in press, Chapter 2) is relevant here. We examined whether the propensity to show cumulative culture is dependent on both context and cognitive ability and proposed that humans may show cumulative culture in a greater range of contexts than nonhumans

due to enhanced cognitive abilities. We investigated whether children of different ages (3-6 years-old), and thus cognitive capacities, were able to improve on information provided by one model (showing the potential for cumulative culture) under two different task contexts. One context was more cognitively challenging and included considerable memory demands, and the other was absent of such demands. We found support for our hypothesis; in the less challenging condition children of all ages were able to improve on the model's score and show the potential for cumulative culture, whereas in the condition which taxed memory only the 6-year-olds showed this potential. This study focussed on an ecologically valid task context in which using information from *one* model presented memory demands. However, using information from multiple models may present similar, and likely greater, demands. Useful information from multiple models may be separated in time and space from a potential social learner's own attempt and exploiting this information is therefore likely to involve increased cognitive load. For example, potentially beneficial information from multiple models may be observed at different time points and remembering and storing this information for later use would be cognitively demanding on memory. Furthermore, information provided by multiple models may contain a range of elements, some which the information user wishes to utilise and others they wish to discount. Holding in mind these different elements, whilst deciding how they are best utilised, may involve working memory (defined as both the storage and manipulation of information, Best & Miller, 2010; Cowan, 2008; Diamond, 2013; Garon et al., 2008).

Alternatively, it may be challenging for a learner to recognise the value of information from another individual in relation to oneself (Blakey et al., 2020). For example, another individual could be interacting with an object equivalent to one currently available in the learner's own immediate vicinity. Using this information will likely involve translating what they have observed to their own bodily frame of reference. This probably requires skills in recognition and mapping of correspondences (DeLoache, 1989, 1991, 2000), and mental translation and rotation (Frick et al., 2013; Levine et al., 1999; Shepard & Metzler, 1971). An example of this would be watching an individual climb a tree to retrieve a coconut and appreciating that one can climb a different tree and retrieve a different coconut. Both temporal and spatial displacement of information therefore generate increased cognitive load for a potential information user. Moreover, we would expect this cognitive load to be further increased if the information to be remembered, mapped or translated comes from more than one source. Thus, effectively utilising information from multiple models is likely to require additional cognitive load. It follows that a learner with a more limited capacity for domain-general cognitive processing would be less able to exploit information presented by multiple models; we expect this to include young children and nonhumans.

Furthermore, if a learner is to overcome the above cognitive challenges to the extent that they can adequately utilise information from multiple models in many real-world examples, metacognitive storage strategies unlikely to be possessed by nonhumans or young children may be needed (Carruthers, 2013). This may include the ability to: deploy attention or resist interference, rehearse items held in memory through inner speech, and/or extend working memory capacity through the use of external storage systems such as writing (Wynn & Coolidge, 2011). We would therefore expect adults to outperform children, and for children to become better at integrating information from multiple models with increasing age. Nonhumans would be expected to find this extremely difficult, or perhaps impossible, depending on the presentation of the information. This theory is supported by the results of our previous work (Wilks et al., in press, Chapter 2) which found that young children were less able to improve on the performance of one model (i.e. show the potential for cumulative culture) if the information needed to be held in working memory rather than remaining continuously present. However, research exploring children's ability to integrate information from multiple models is in its infancy, and not directly comparable to the adult experimental work. Nevertheless, we will discuss two studies which are a relevant starting point.

Subiaul et al. (2015) have studied 3-5-year-olds' ability to open a two-compartment box when provided with adequate demonstrations from one model (opening both compartments), two models (each opening one compartment) or in an individual learning condition with no demonstrations. They found that children were more likely to successfully open both compartments in the model conditions compared to the individual learning condition, but that there was no difference in performance between the one and two model conditions. This is not surprising, because the actual information provided in the one and two model conditions was the same – the demonstration from the two models as a whole was identical to that provided by one model. This study therefore showed that the presence of two models did not present a problem in terms of using the available information. However, unlike in the adult experimental work, in this study using the information from two models does not require any of the realistic constraints which might operate when integrating information from multiple models in the real world, such as separation of the information in time (Fay et al., 2019) and/or space (Kempe & Mesoudi, 2014; Muthukrishna et al., 2014). Additionally, this study did not require children to be selective regarding the information they used from each model – each provided exactly half of that in the full, one model, demonstration. Thus, simply summing the information across the two models would result in perfect performance. In real-world cases of cumulative culture, learners are more likely to be exposed to multiple, imperfect demonstrations which together may contain all the information

needed, but which also contain potentially distracting information about other behaviours not linked to success. The selective extraction of this relevant information, and selective inhibition of any redundant or ineffective elements, probably brings with it significant cognitive challenges that were not captured within the design of Subiaul et al. (2015).

An experiment by Buchsbaum et al. (2011) has demonstrated that 3-5-year-olds do have the ability to use information selectively, although this study did not look specifically at learning from multiple models. Children were shown five different action sequences (which caused a toy to play music) and were able to integrate this information to eliminate actions which were causally irrelevant. However, this study did not require children to integrate *successful* elements across different demonstrations. Nevertheless, it does demonstrate that children have the ability to integrate information selectively across multiple demonstrations.

The developmental literature outlined above demonstrates that children can use information from multiple models in principle, and that they can integrate simple information selectively. However, there is currently no evidence that they can use information from multiple models to the extent that an improved solution (improved trait variant) is introduced into an experimental population. Consequently, further work is required before we can draw conclusions as to children's ability to use combinatorial mechanisms to generate examples of cumulative culture.

Building on the results of Subiaul et al. (2015), and the aforementioned adult studies (Kempe & Mesoudi, 2014; Muthukrishna et al., 2014), we explored children's ability to use information from multiple models to create an improved variant. Children were presented with information from two models and we were interested in whether they could extract relevant information, and eliminate that which was ineffective, to perform more highly than the highest performing single model in this two-model population. Furthermore, we investigated how this ability changed during development and according to different cognitive constraints presented in four conditions. This would enable us to draw conclusions as to the potential impact and importance of cognitive demands on combinatorial social learning and cumulative culture, and to analyse whether human-unique cognitive processes may be necessary for this kind of information use.

We tested children across a wide developmental span (ages 5-10). Across our four experimental conditions the information provided by the two models differed in its temporal and/or spatial displacement, which provided varied and realistic cognitive constraints as discussed above. We did not test children below age 5 because our previous work using a similar task, which included a condition taxing working memory, had shown that until age 6 children found using information from even just one model extremely difficult (Wilks et al., in press, Chapter 2). We thus predicted that children under age 5 would find certain conditions

in our current task even more challenging, but that testing children from age 5 to 10 should provide sufficient scope to document age-related effects on task performance.

The information from the two models was presented to children in the form of a searching task presented on a touchscreen computer, in which both models individually searched a space for hidden treasure. In our aforementioned study, we had successfully used a similar paradigm (although not computerised) with a single model and found it to be understandable and informative (Wilks et al., in press). In the current study, the rewards (“treasure” – gold coins) were always hidden in three of nine chests presented as 3x3 arrays on each of 24 trials. Each model made three selections from each, nine-chest array, so the maximum number of reward locations which could theoretically be revealed, both by each model and collectively across the two models on each array, was three. We placed constraints on the selections made by both models; these conformed to six different demonstration types (Table 3.1), each of which was presented once within each of four conditions (described below). These demonstration types varied in relation to the highest score achieved by a single model in the pair (equivalent to the number of rewards found by this model – 1 or 2), the mean score across the two models, and the total number of rewards found (out of 3) by both models combined (see Table 3.1 and pages 78-79 for further details).

Table 3.1

Summary of the Six Demonstration Types Denoting the Models' Selections (Each Presented Once Within the Increased Information, Temporal, Spatial and Temporal-Spatial Conditions)

Highest single model score	Demo type	Number rewarded (R) & unrewarded (U) selections – model 1	Number rewarded (R) & unrewarded (U) selections – model 2	Mean score across the two models	Expected chance score	Total number of rewarded locations revealed	Total number of unrewarded locations revealed	Number of rewarded locations selected by both models (repeated)	Number of unrewarded locations selected by both models (repeated)	Number of rewarded locations selected by one model only	Number of unrewarded locations selected by one model only	Number of locations with unknown contents following models' selections
1	3	R UU	R UU	1.00	1.00	2	3	0	1	2	2	4 (1R, 3U)
1	5	R UU	UUU	0.50	1.00	1	3	0	2	1	1	5 (2R, 3U)
1	6	R UU	R UU	1.00	1.00	1	3	1	1	0	2	5 (2R, 3U)
1	Mean 3,5,6			0.83	1.00							
2	1	RR U	R UU	1.50	1.00	3	2	0	1	3	1	4 (4U)
2	2	RR U	RR U	2.00	1.00	3	2	1	0	2	2	4 (4U)
2	4	RR U	R UU	1.50	1.00	2	2	1	1	1	1	5 (1R, 4U)
2	Mean 1,2,4			1.67	1.00							

We expected the least challenging of our four conditions to be our “Increased Information” condition in which the information provided by the models was not displaced either temporally or spatially. That is, both models sequentially searched the array which the participant themselves would later search. In addition, both models’ selections remained visible in the search space whilst the participant made their own selections. This condition acted as our control because the information provided was effectively the same (at least in terms of the cognitive challenges involved) as if it had been provided by a single model. However, the condition also controlled for any distraction or facilitation caused by the presence of two models. In contrast, our “Temporal” displacement condition presented a greater cognitive challenge. Although the models once again made their selections from the same array, the locations selected by each model were concealed again, immediately after the model’s third selection. Therefore, in order to fully integrate the information, the locations and reward value (rewarded/unrewarded) needed to be held in memory. We also included a “Spatial” displacement condition in which the first model made their selections from an array on the top left of the screen, the second model from an identical array on the top right, and the child from another such identical array on the bottom middle of the screen. The rewards were hidden in the same positions across the three arrays, but we were interested in whether children could integrate useful information from the top left and right arrays. As in the Increased Information condition, the reward locations selected by the models remained visible throughout the child’s selections. Our “Temporal-Spatial” displacement condition was expected to be the most cognitively challenging because it combined the temporal and spatial elements from the Temporal and Spatial conditions respectively. That is, the models made their selections on two arrays, as in the Spatial condition, but the revealed locations were concealed again immediately after each model made their third selection, as in the Temporal condition. To utilise the information fully therefore required both an ability to integrate information from two spatially distinct locations and the capacity to hold this in one’s working memory.

Our expectation was that children would utilise the information from the two models across a wider range of conditions as their ability to overcome additional cognitive demands increased with age. Moreover, we expected that children’s ability to repeat the rewarded information, and not repeat the unrewarded information, provided by the models would differ depending on age and condition i.e. the different working memory and spatial integration constraints presented.

Measuring the Potential for Cumulative Culture

In order to investigate children's ability to produce an improved trait variant, and hence potential for cumulative culture, we grouped according to chronological age, using one-year bandings, and split the data further depending on condition (Increased Information, Temporal, Spatial or Temporal-Spatial) and whether the highest single model score in a given trial was 1 (henceforth model-score-1 trials) or 2 (model-score-2 trials). We also calculated the mean child score for each age group, condition, and model-score groups (1 or 2). For the model-score-2 trials only, we planned to analyse whether these mean scores were significantly greater than the highest single model score of 2. Outperforming the highest observed single model score was analogous to outperformance of the best available information within a population, an improved variant. Additionally, again for the model-score-2 trials only, we were interested in whether the mean child scores were significantly greater than the mean score across the two models. Outperforming the mean score of the two models could be used as a proxy for outperformance of a randomly selected single model. Linking back again to our interest in cumulative culture, this measure would give an indication of whether children could outperform the previous generation's "typical" score. This is important because it would show whether, *on average*, later generations would accumulate benefits relative to their predecessors. We also measured whether children were scoring significantly above chance level (i.e. the score expected to arise from naive exploration in the absence of any information) for both the model-score-1 and 2 trials. A chance level score was 1 because each child selected three out of nine chests on each array and there was a total of three rewarded chests within each array. This comparison with chance would serve to test whether the children were indeed using some of the information available to them in the demonstration (even if not optimally). This is because scores not significantly different from chance could not be distinguished from a pattern of random selections, uninfluenced by the demonstration information. In the model-score-1 trials this chance-level of 1 was equivalent to the highest single model score. Therefore, for the model-score-1 trials, scoring significantly higher than the highest single model score was effectively captured by this same comparison. Furthermore, in the model-score-1 trials, the mean score across the two models (0.83) was lower than the chance-level therefore outperformance of the mean did not provide us with meaningful information, over and above the comparison to chance, for the same reason. The benchmarks of outperforming the highest single model score and mean score, as used for the model-score-2 trials, were therefore not meaningful for the model-score-1 trials.

Materials and Methods

Participants

169 children were recruited from a primary school in Stirling, Glasgow Science Centre, and a science festival attended by members of the public at the University of Stirling. Five children (all male) were excluded due to: missing date of birth from the consent form (age 8), failure to fully comply with task instructions (age 10), a recognised developmental delay ($n=2$, aged 6 & 7) or because they were found to have participated previously (age 6). The final sample consisted of 164 children aged 5 to 10 ($M= 7$ years, 10 months; range= 5 years, 0 months – 10 years, 11 months; $SD= 1$ year, 8 months; 79 female); between 20 and 31 children in each age group.

Ethics statement: This research was approved by the University of Stirling, General University Ethics Panel (reference: GUEP599). Written, informed consent was obtained from the parent or guardian of all children prior to their participation. Children were asked if they would like to participate, were continuously monitored for assent and were rewarded with a sticker regardless of task completion.

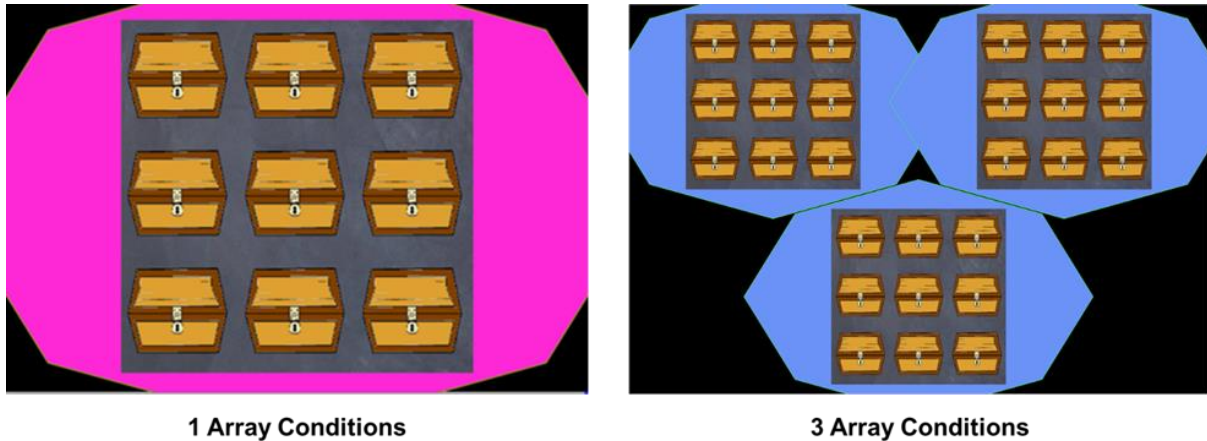
Experiment

The task was presented to children on either a touchscreen laptop or tablet running Windows 10 and was written and run in PsychoPy, version 1.84.2 (Peirce et al., 2019). Task responses were automatically written into a csv file and any verbal comments children made were recorded on paper by a research assistant.

The goal of the task was to find as many pieces of treasure (gold coins) as possible following watching two social models attempt the same task. Each of 24 experimental trials (plus four practice trials) contained either one (Increased Information and Temporal displacement conditions, Figure 3.1, left) or three (Spatial displacement and Temporal-Spatial displacement conditions, Figure 3.1, right) 3x3 array(s) of nine treasure chests, three chests of which were rewarded and six unrewarded. Rewarded chests revealed a gold coin, accompanied by the sound of money, if selected, and unrewarded chests a red cross and a beep. Participants took part in four experimental conditions, which varied in terms of whether the information provided by the two models was separated from their own attempt, either in time or space (Increased Information, Temporal displacement, Spatial displacement and Temporal-Spatial displacement, see introduction). As stated in our introduction, we placed constraints on the selections made by both models which conformed to six different demonstration types (further details in Table 3.1 and pages 78-79).

Figure 3.1

Example Task Display Containing One (Left Image) or Three (Right Image) Arrays as in the Increased Information/Temporal and Spatial/Temporal-Spatial Conditions Respectively



To determine the order of the conditions, we randomly selected from the 24 different permutations of the four conditions (Increased Information, Temporal, Spatial and Temporal-Spatial) for each participant. The two social models were a cartoon parrot “Pirate Parrot” and octopus “Pirate Octopus”. Between participants, we counterbalanced which model made their selections first. Therefore, the model choosing first remained consistent for a participant throughout the task. The order of the six demonstration types was randomised for each participant within each condition. Within a demonstration type, we also randomised whether the parrot or the octopus performed the selections designated for model 1 or model 2 and the actual locations selected within the array.

Demonstration Types. Across the six demonstration types (1-6, Table 3.1), the highest single model score over the two models on any given trial was either 1 or 2 (three of each type); the mean score across the two models varied between 0.5 and 2.0; the total number of rewards revealed across the two models combined was either one, two or three (two of each type); and the number of rewarded selections repeated by each model within an array was either 1 or 0 (three of each type). We did not include a demonstration with zero rewards because we wanted to ensure that children were receiving some information that it would be of benefit to copy on each trial. We did not include demonstration types in which any one model found all three rewards because we were interested in children’s ability to use information from the two models to find more rewards than the highest performing

model. We were also interested in how well children made use of social information about locations of rewards (arising from correct choices) and locations to be avoided (arising from incorrect choices).

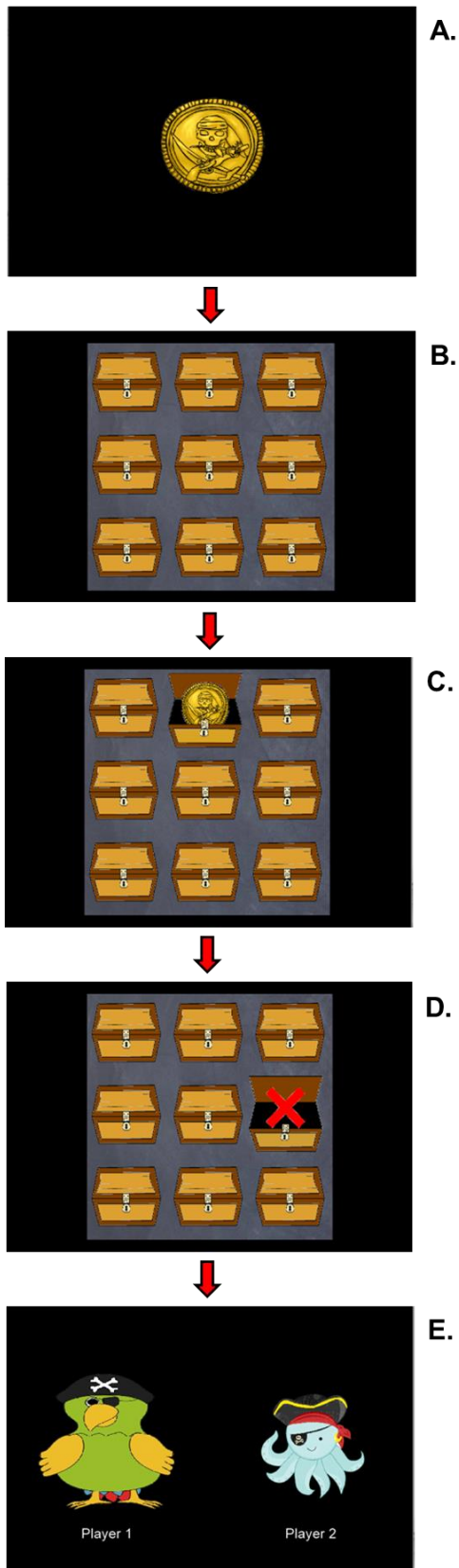
Procedure

In the school, testing took place in a quiet area adjacent to the classroom. At Glasgow Science Centre and a science festival held at the University of Stirling, testing was conducted in a public space, separated from the main museum/festival space by a desk or room partition. At the science centre and festival only, less confident children were accompanied in the testing area by a parent or guardian who was instructed not to provide the child with any assistance relevant to the task. In each location the laptop or tablet running the task was positioned on a table-top, with the child sat in front of this at a comfortable distance and the experimenter sat next to the child. A verbal script was used by the experimenter (see Appendix “Experimenter Verbal Scripts”) and total testing duration was 15-20 minutes per child.

Introduction to Task. Children were asked if they would like to play a game in which the goal was to search for treasure. They were first shown a series of on-screen images (Figure 3.2), supported by verbal instruction and explanation from the experimenter. The gold coins, arrays of treasure chests, notion of rewarded/unrewarded chests and the models were all introduced. Children were informed that the models would also be looking for treasure and that they should watch them carefully before taking their turn. The goal of the experiment, and the fact that they should try to find all three pieces of treasure in each array, were reinforced several times throughout the introduction.

Figure 3.2

Images Displayed to Children During the Task Introduction



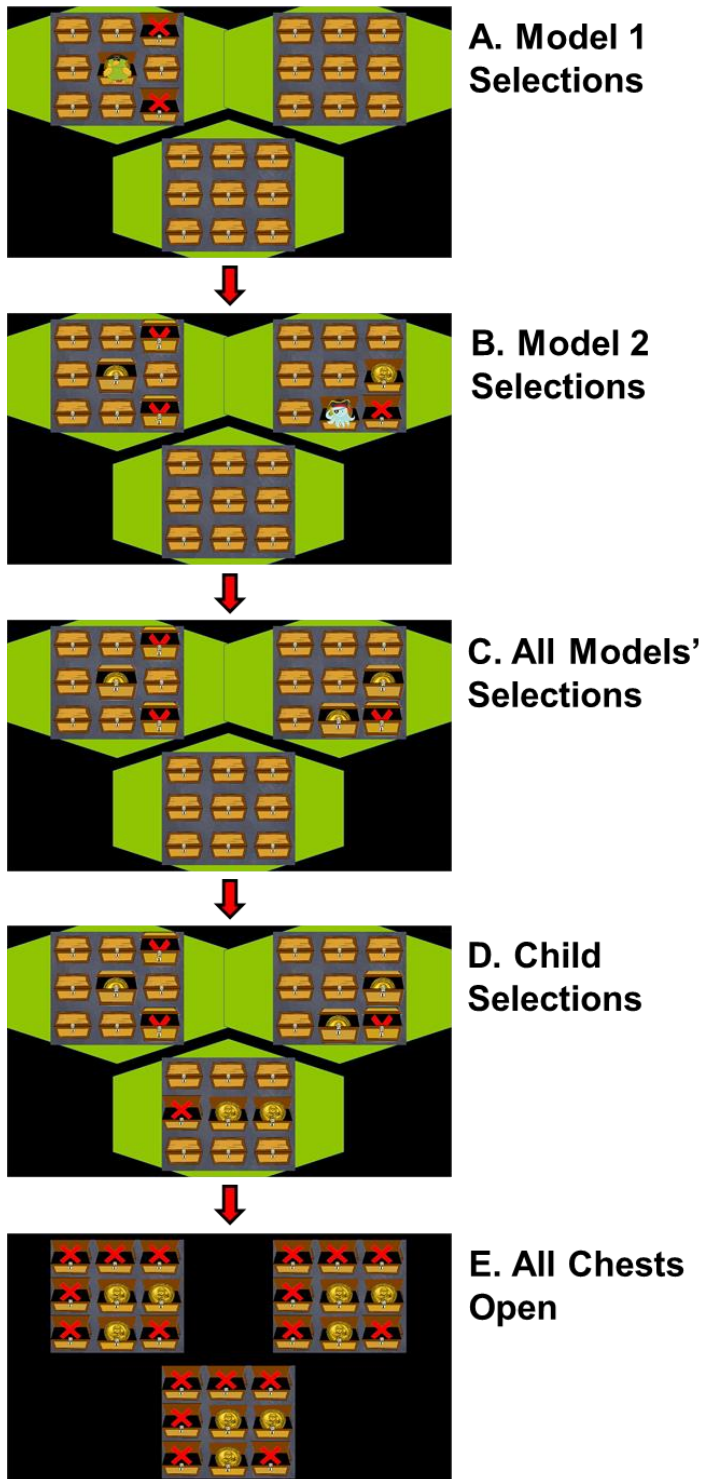
Practice Trials (x1 Per Condition). Following the introduction, a practice trial corresponding to the first assigned condition was conducted prior to the six experimental trials for that condition. The six experimental trials covered each of the six demonstration types (Table 3.1). The remaining three conditions were run in the same way, with a practice trial followed by the six experimental trials.

A practice trial began with the presentation of a single array set back slightly from the forefront, in the middle of the screen. This array then moved forward for the single array conditions (Increased Information and Temporal) or gave the illusion that it was splitting into three for the three array conditions (Temporal-Spatial and Spatial). In the latter case the top half of the screen contained one array on the left and one on the right, and the bottom half one array in the middle (Figure 3.1, right image). However, the rewards were in the same location in each of the three arrays. The first of the two models then selected three chests (see Figure 3.3A for an example for the Spatial condition) and upon each selection a chest opened to reveal either a coin (rewarded) or a red cross (unrewarded) and accompanying sounds. Immediately following the third selection chests either closed completely, hiding their content, or partially, meaning that their content remained visible. In the conditions with three arrays this first model's selection was made on the top, left array (Figure 3.3A). The second model then proceeded to select three chests in the same manner as the first. In the single-array conditions this was on the same array as the first model and in the three array conditions the top, right array (Figure 3.3B). All selected chests then either closed completely, or partially (Figure 3.3C). Next, the child was prompted to search for all three pieces of treasure, and it was made clear that they could select a chest even if one, or both, of the models had already selected it. We wanted to ensure (in the Increased Information and Spatial conditions especially) children understood that they could find treasure in the same locations as the models and that a model selecting a correct chest did not mean that the treasure had been removed from this chest. Children made their selections by touching three of the nine chests in turn on either the same array as the two models (one array conditions), or the array in the bottom, middle of the screen (three array conditions), Figure 3.3D. As occurred for the models' selections, the chests initially opened to reveal their contents as they were chosen, but after the third selection the three chests remained open and the remaining six, unselected chests, also opened so that the contents of all nine chests were visible (Figure 3.3E). The experimenter then stated "Look – that's where the treasure was" and pointed to the rewarded chests in turn in order to reinforce that exactly three of the nine chests were rewarded on every trial. The child was asked: "How many pieces of treasure were there?" This question assessed understanding and if a child answered

correctly or incorrectly their response was reinforced or corrected respectively.

Figure 3.3

Diagram of Testing Procedure: Model 1 Selects Three Chests (A), Model 2 Selects Three Chests (B), Screen Display Following All Models' Selections (C), Child Selects Three Chests (D) and All the Chests Open to Reveal the Contents (E)



Note. This diagram depicts the Spatial condition as an example. The exact images displayed on screen varied depending on the condition e.g. for the Temporal and Temporal-Spatial conditions all chests in image C would be closed.

Experimental Trials (x 6 Per Condition). The six experimental trials (spanning the six different demonstration types – Table 3.1) followed the practice trial for that condition and were almost identical except that there was less intervention from the experimenter. The participant was not reminded that they were searching for three pieces of treasure and was not asked how many pieces of treasure were revealed following their selections, although the experimenter still pointed to each of the three pieces of treasure when all chests opened at the end of the trial. The only other difference was that the single or three arrays (Figure 3.1) appeared immediately at the start of the trial and did not need to be moved forward (single array condition) or give the illusion of splitting into three (three array condition).

Ending of Experiment. Following completion of all four conditions (each consisting of a practice trial followed by six experimental trials) a black screen was presented. A large, gold coin superimposed onto a chest appeared in the middle of the screen and a number above the chest continued to increase incrementally until it represented the total number of rewards found by the participant – a maximum of 72 (3 per array x 24 trials, the rewards from practice trials were not included).

Results

Firstly, we were interested in how overall use of the information provided by the two models (rewarded and unrewarded) differed with age according to the constraints presented by our four conditions (see our analysis of score and repeating). Secondly, we were interested in whether children aged 5-10 were able to make use of the information available from the two models across our four conditions to the extent that they outperformed the highest scoring single model (see “Potential for Cumulative Culture”). This question was of particular interest as a result of our motivation to understand the potential for cumulative culture. In this context, outperforming the highest scoring model could be likened to producing an improved trait variant, which goes beyond the achievements of the previous “generation”.

p -values $< .05$ were taken as statistically significant across all analyses. All generalised linear mixed effects models (GLMM) were carried out with the log link (count data, family = poisson) and the lme4 package (lmer function) using R (R Core Team, 2018). Our default choice for the random effects structure for each model included by-participant

random slopes for variables which varied within participant, following Barr et al. (2013), to keep random effects structures “maximal” where possible. Where the maximal model resulted in non-convergent or singular fit models, random slopes were removed followed by random intercepts where necessary until a convergent, non-singular model was obtained.

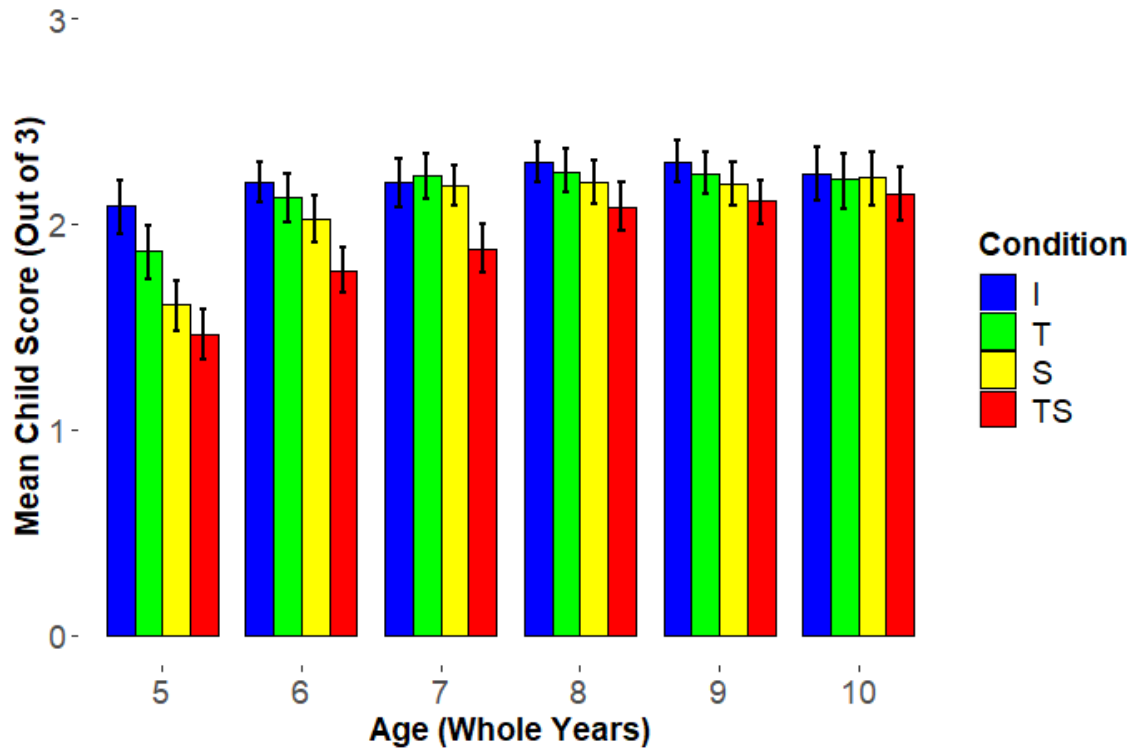
Score

We measured overall information use across our four conditions, in ages 5-10, according to child score (out of 3) per array. Children were given a point for each piece of treasure found on an experimental trial therefore the total score on each array was 0, 1, 2 or 3. This measure was useful in order to provide an overview of any effects of age and condition on overall child performance. However, it was a rather crude measure because it did not capture how children responded to demonstrations of different reward value, and to rewarded and unrewarded model selections. These were captured in our other analyses.

Score – Effects of Age and Condition. We performed a GLMM with score (out of 3) as the dependent variable; and age and condition, and the interaction of these variables, as fixed effects. Condition was dummy coded, with each condition compared to the reference, Increased Information, condition. We also included a random intercept for the total number of rewards revealed in a trial (1, 2 or 3 across the two models). There was a main effect of condition: lower scores (out of 3) in the Spatial ($b = -0.077$, $SE = 0.031$, $Z = -2.505$, $p < .050$) and the Temporal-Spatial ($b = -0.164$, $SE = 0.032$, $Z = -5.200$, $p < .001$) conditions compared to the reference, Increased Information, condition (Figure 3.4). In the Temporal condition, score was not significantly different from that in the reference condition ($b = -0.032$, $SE = 0.030$, $Z = -1.047$, $p = .295$). There was no main effect of age ($b = 0.043$, $SE = 0.035$, $Z = 1.233$, $p = .218$) but there was a significant interaction effect of age on the Temporal-Spatial ($b = 0.154$, $SE = 0.052$, $Z = 2.983$, $p < .010$) condition, showing that age affected score differently in this condition, compared to the reference condition. There was no such interaction effect for the Temporal condition ($b = 0.036$, $SE = 0.050$, $Z = 0.726$, $p = .468$) but this was approaching significance for the Spatial condition ($b = 0.098$, $SE = 0.050$, $Z = 1.942$, $p = .052$). Post hoc comparisons between all four conditions were carried out using the EMMEANS package in R. These revealed that score was significantly different for the Temporal-Spatial condition compared to the Temporal ($SE = 0.036$, $p < .001$) and Spatial ($SE = 0.035$, $p < .010$) conditions in addition to the Increased Information condition, as reported above. There was also a significant difference between the Spatial and Increased Information condition, again as reported above. There were no other significant differences between the conditions (all $p > .100$).

Figure 3.4

Mean Child Score/3 by Condition and Age (Whole Years)



Note. The conditions are abbreviated as follows: I = Increased Information, T = Temporal, S = Spatial and TS = Temporal-Spatial. Error bars are 95% confidence intervals.

Score – Effects of Age by Condition. We ran four exploratory GLMMs (one for each condition), with a view to further investigating the differing effects of age on score in each of the four conditions (i.e. the interaction between age and condition). We split the data by condition and removed this from the original model. The current models were therefore identical to the first model (p. 82) except that they had a fixed effect of age only. We found a main effect of age in the Temporal ($b = 0.079$, $SE = 0.036$, $Z = 2.233$, $p = .026$), Spatial ($b = 0.141$, $SE = 0.036$, $Z = 3.882$, $p < .001$) and Temporal-Spatial ($b = 0.197$, $SE = 0.038$, $Z = 5.193$, $p < .001$) conditions but not in the Increased Information condition ($b = 0.043$, $SE = 0.035$, $Z = 1.233$, $p = .218$). This showed that score increased with increasing age in all conditions except the Increased Information (Figure 3.4).

Score – Effects of Condition by Age. We also ran six more exploratory GLMMs (one for each age group), with a view to further investigating the differing effects of condition on score by age. We split the data by age and removed age from the original model therefore the current models were identical to the first model (p. 82) except that they had a fixed effect of condition only. We found main effects of conditions Temporal-Spatial ($b = -0.355$, $SE = 0.083$, $Z = -4.277$, $p < .001$) and Spatial ($b = -0.262$, $SE = 0.081$, $Z = -3.243$, $p = .001$) in our model for children aged 5: lower score in these conditions compared to the Increased Information condition (Figure 3.4). For children aged 6 ($b = -0.217$, $SE = 0.074$, $Z = -2.936$, $p = .003$) and 7 ($b = -0.160$, $SE = 0.079$, $Z = -2.020$, $p = .043$) there was a main effect of the Temporal-Spatial condition only: lower score in this condition compared to the Increased Information. There were no main effects for the remaining conditions in children aged 5, 6 and 7; for 8-10-year-olds there were no main effects for any condition (all $p > .100$): score did not significantly differ from that in the Increased Information condition.

Repeating

Repeating the models' rewarded selections, and not repeating the models' unrewarded selections, are both correct strategies but pose quite different demands. While the former requires one to remember a rewarded selection and repeat it, the latter requires one to remember an unrewarded selection and avoid repeating it. If the social information from the two models was being integrated effectively, we would expect high and low levels of repeating following rewarded and unrewarded model selections respectively. We therefore analysed the total number of rewarded and unrewarded selections repeated ("rewarded repeats" and "unrewarded repeats") on each trial, for each participant, and how this differed according to age (thousands of days, centred) and condition (Increased Information, Temporal, Spatial, and Temporal-Spatial) in the below GLMMs.

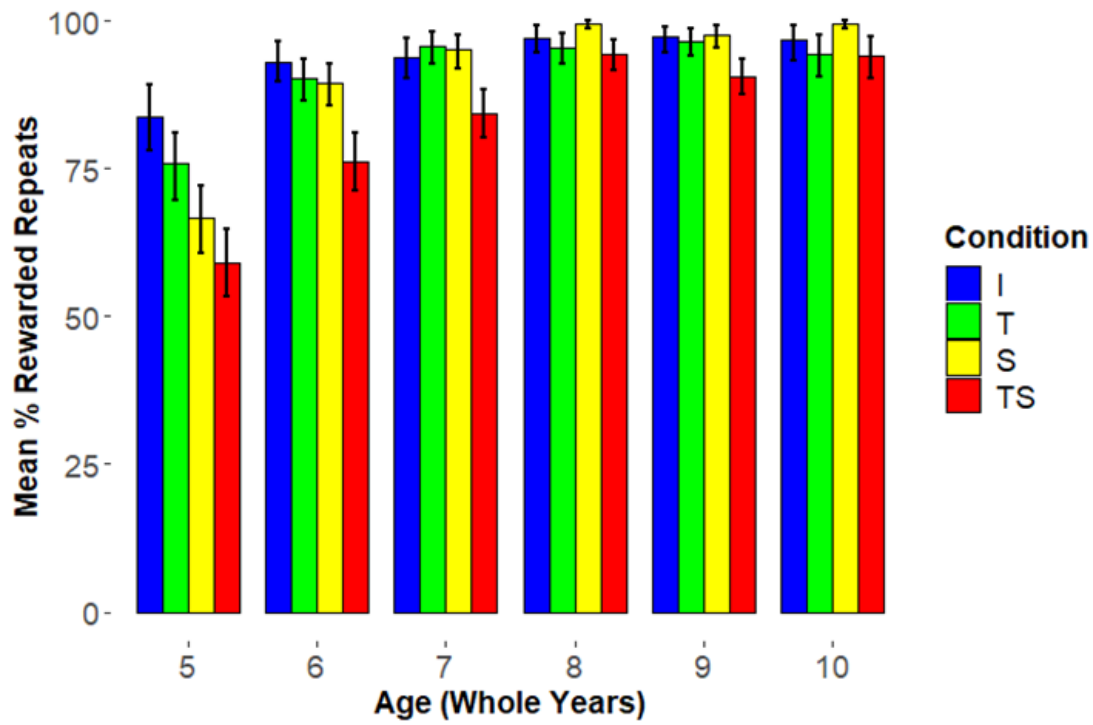
Rewarded Repeats – Effects of Age and Condition. We performed a GLMM with the number of repeats of rewarded selections as the dependent variable; and age and condition, and the interaction of these variables, as fixed effects. Condition was dummy coded, with each condition compared to the reference, Increased Information, condition. We also included a random intercept for the total number of rewards revealed in a trial (1, 2 or 3 across the two models). There was a main effect of condition ($b = -0.154$, $SE = 0.034$, $Z = -4.498$, $p < .001$): fewer repeats of rewarded selections in the Temporal-Spatial condition compared to the reference, Increased Information, condition (Figure 3.5A, Table 3.2). The Spatial and Temporal conditions were not significantly different from the reference condition (all $p > .100$). There was no main effect of age ($b = 0.065$, $SE = 0.038$, $Z = 1.718$, $p = .086$)

but there was a significant interaction effect of age on the Temporal-Spatial ($b = 0.166$, $SE = 0.056$, $Z = 2.965$, $p < .001$) and Spatial ($b = 0.108$, $SE = 0.054$, $Z = 1.983$, $p = .047$) conditions, showing that age affected the number of repeats differently in these conditions, compared to the reference condition. There was no such interaction effect for the Temporal condition ($b = 0.034$, $SE = 0.054$, $Z = 0.622$, $p = .534$). These interaction effects mirrored those of age on condition in the analysis of score on page 85 (except that here the interaction between age and the difference between the Spatial and reference conditions reached significance). Therefore, to avoid repetition, we do not report further models to investigate these interaction effects here (as we did for score), but instead in the Appendix. Post hoc comparisons between all four conditions were carried out using the EMMEANS package in R. These revealed that the number of rewarded repeats was significantly different for the Temporal-Spatial condition compared to the Temporal ($SE = 0.031$, $p < .001$) and Spatial ($SE = 0.031$, $p = .001$) conditions in addition to the Increased Information condition, as reported above. There were no other significant differences between the conditions (all $p > .100$).

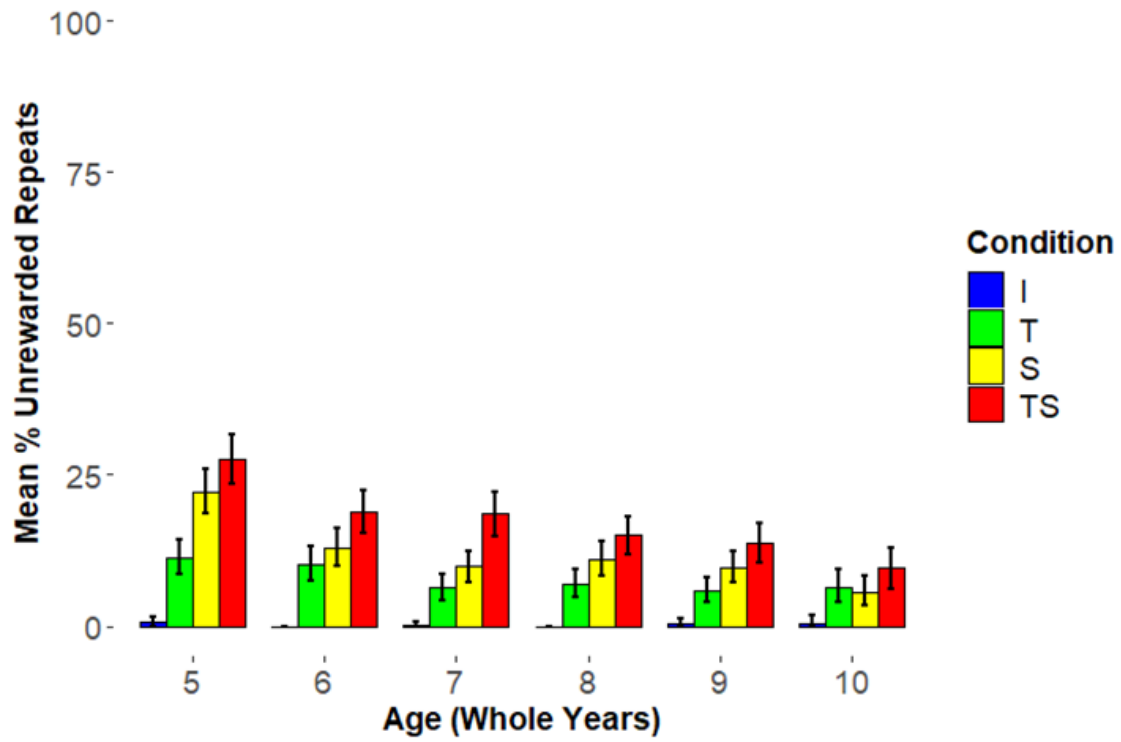
Figure 3.5

Mean Percentage of Rewarded (Panel A) and Unrewarded (Panel B) Model Responses Which Were Repeated by Condition and Age (Whole Years)

A.



B.



Note. The conditions are abbreviated as follows: I = Increased Information, T = Temporal, S = Spatial and TS = Temporal-Spatial. Error bars are 95% confidence intervals.

Table 3.2

Mean Percentage of Rewarded and Unrewarded Model Selections Repeated (and Standard Deviation) for Ages 5-10 (Whole Years) in the Increased Information (I), Temporal (T), Spatial (S) and Temporal-Spatial (TS) Conditions

Age (whole years)	Condition	Mean percentage rewarded repeats	SD (rewarded repeats)	Mean percentage unrewarded repeats	SD (unrewarded repeats)
5	I	83.730	35.646	0.595	5.439
6	I	93.100	22.698	0.000	0.000
7	I	93.910	22.085	0.214	2.669
8	I	97.123	15.900	0.000	0.000
9	I	97.222	14.798	0.448	4.395
10	I	96.667	18.026	0.556	6.086
5	T	75.893	36.147	11.310	19.671
6	T	90.233	24.906	10.215	20.582
7	T	95.727	17.948	6.410	14.475
8	T	95.437	17.656	6.944	15.414
9	T	96.595	15.608	5.824	14.410
10	T	94.444	19.842	6.389	16.481
5	S	66.567	39.436	22.222	23.758
6	S	89.427	24.481	12.903	21.029
7	S	95.085	17.982	9.829	17.747
8	S	99.405	4.428	11.111	19.785
9	S	97.491	13.071	9.677	18.063
10	S	99.444	4.285	5.694	14.085
5	TS	58.929	38.986	27.480	27.306
6	TS	76.254	32.638	18.907	24.188
7	TS	84.295	26.790	18.483	23.915

8	TS	94.444	17.933	14.980	21.783
9	TS	90.681	21.289	13.620	20.931
10	TS	94.167	19.404	9.583	18.418

Unrewarded Repeats – Effects of Age and Condition. We performed a GLMM which was identical to the model on pages 87-88 except that we used the number of unrewarded repeats as the dependent variable. There were main effects for each condition (all $p < .001$): more unrewarded repeats in the Temporal, Spatial and Temporal-Spatial conditions compared to the reference, Increased Information, condition (Figure 3.5B, Table 3.2). There was no main effect of age ($b = 0.414$, $SE = 0.628$, $Z = 0.659$, $p = .510$) and no interaction effects between age and the Temporal, Spatial or Temporal-Spatial conditions (all $p > .100$). Post hoc comparisons between all four conditions were carried out using the EMMEANS package in R. These revealed that the number of unrewarded repeats was significantly different for all four conditions compared to all other conditions (all $p < .001$).

Potential for Cumulative Culture

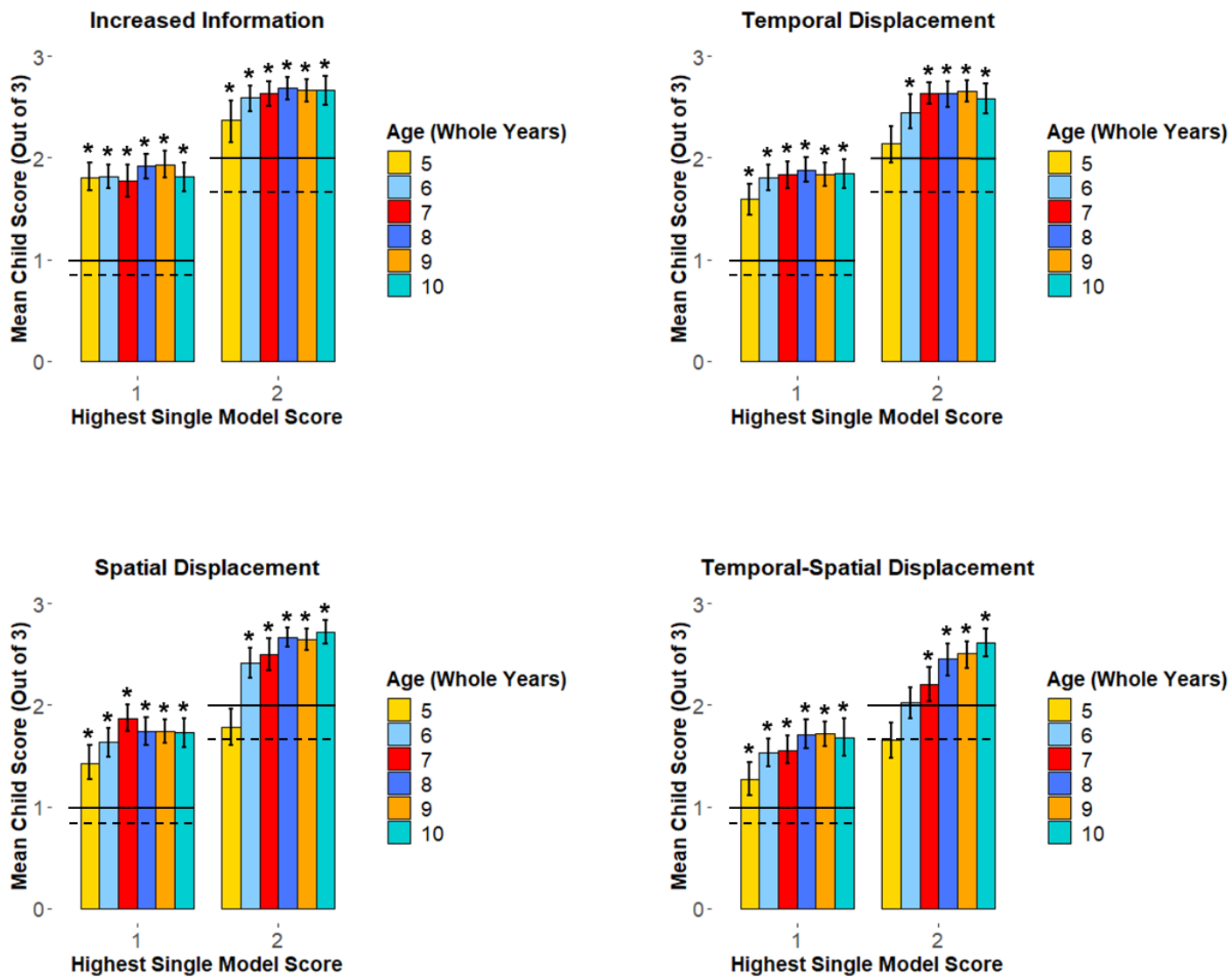
In order to measure children's potential for cumulative culture, based on the different patterns of performance outlined in our introduction, we measured whether the mean child score (out of 3) on the model-score-2 trials (for each age group and condition) adhered to the criteria outlined in four levels. Mean score needed to be as follows to achieve each level: Level 0: at or below the chance-level of 1; Level 1: significantly greater than chance-level; Level 2: significantly greater than both chance-level and the mean score across the two models; Level 3: significantly greater than the highest single model score.

Performance for the model-score-2 trials differed according to age and condition (Figure 3.6, Table S.4). Children performed most proficiently in the Increased Information (control) condition in which all age groups scored significantly higher than the highest single model score of 2, hence achieved Level 3 (ages 6-10, $p < .001$; age 5, $p = .011$; Figure 3.7). However, in the Temporal condition the level achieved differed with age. Ages 6-10 achieved Level 3 (all p values $< .001$). Children aged 5 did not outperform the highest single model score ($p = .145$) but achieved a Level 2 because they performed significantly above chance level ($p < .001$) and significantly above the mean score across both models ($p < .001$). This showed that they were performing better than naïve exploration and gaining some performance benefit from the presence of two models. In the Spatial condition ages 6-10 again achieved a Level 3 (all p values $< .001$). Children aged 5 did not outperform the highest single model score ($p = .941$) and although they performed significantly above chance ($p <$

.001) they did not significantly outperform the mean score across the models ($p = .190$) and were therefore given a Level 1. In the Temporal-Spatial condition 5-year-olds again scored a Level 1 i.e. they did not outperform the highest single model score ($p = .941$) or the mean score across the models ($p = .547$) but they did outperform chance ($p < .001$). However, 6-year-olds performed more poorly than in all other conditions, achieving a Level 2 because they did not outperform the highest single model score ($p = .411$) but outperformed both the mean score ($p < .001$) and chance ($p < .001$). Outperformance of the highest single model score, Level 3, was therefore found only in ages 7-10 (ages 8-10, $p < .001$; age 7, $p = .010$).

Figure 3.6

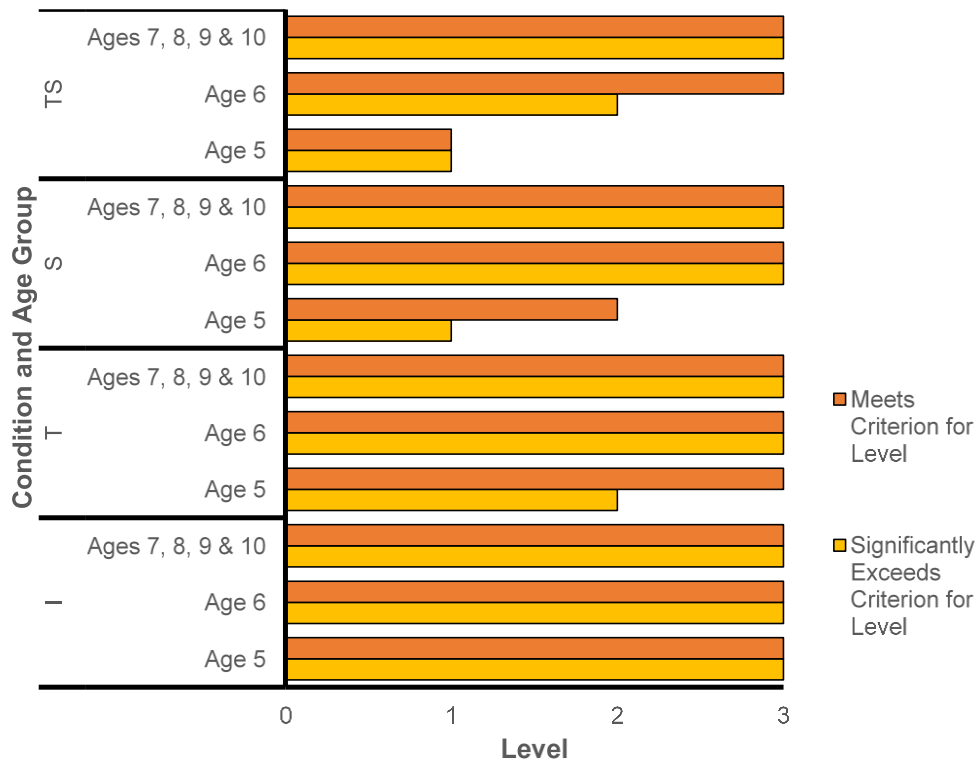
Mean Child Score/3 by Age (Whole Years) and the Highest Single Model Score in a Trial for the Increased Information, Temporal, Spatial and Temporal-Spatial Conditions



Note. The solid line at 1 depicts chance performance: the score expected if children were selecting chests at random and not using the social information. The dashed lines at 0.83 (model-score-1 trials) and 1.67 (model-score-2 trials) depict the mean scores across the two models in the respective trials. The solid line at 2 allows visualisation of whether children scored above 2 when the highest single model score in a trial was 2. An asterisk indicates that the mean child score/3 is significantly above 1 or 2 for the model-score-1 and model-score-2 trials respectively. Error bars are 95% confidence intervals.

Figure 3.7

Potential for Cumulative Culture Performance Classification Level by Condition and Age Group



Note. Situations in which a mean child score/3 was *significantly* higher (yellow bars) and *higher* but not significantly higher (orange bars) than the benchmark outlined in a particular level criterion are described using the terms “Significantly Exceeds Criterion for Level” and “Meets Criterion for Level” respectively.

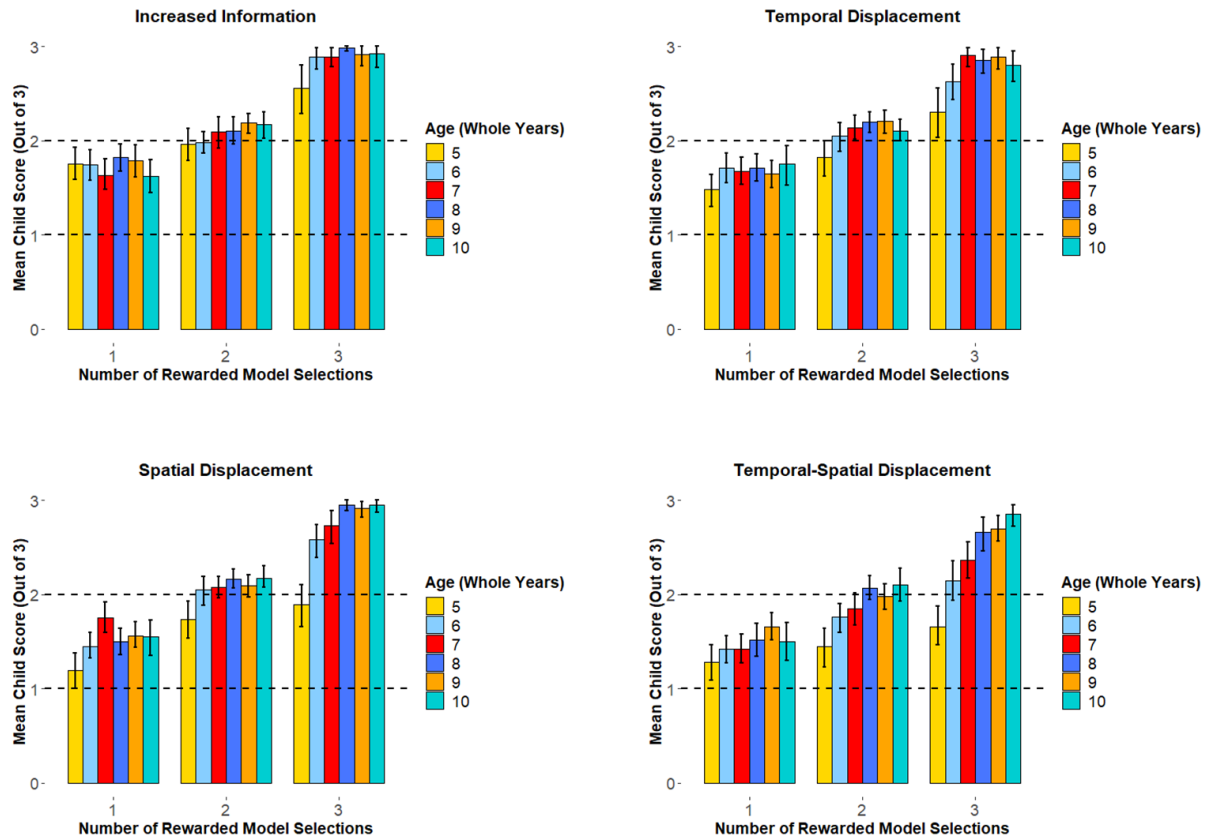
As stated above, for the model-score-1 trials it was only useful to determine if children scored more highly than chance (1.0) and therefore whether they were gaining some benefit from the information provided by the models (equivalent to Level 1). As in the model-score-2 trials, each age group, in each condition, performed significantly above chance-level; ages 6-10, $p < .001$; age 5, $p = .004$ (Figure 3.6, Table S.5).

Across the six demonstration types (outlined on pages 78-79), three were model-score-1 trials (3, 5 and 6, Table 3.1), and three model-score-2 trials (1, 2 and 4, Table 3.1). The three demonstration types present within these model-score-1/model-score-2 groups

were obviously identical in terms of highest single model score but differed in factors such as the total number of rewarded locations revealed and the number of rewarded/unrewarded selections repeated etc. (see Table 3.1 for specific differences). In addition to the above analysis for the model-score-1 and model-score-2 groups, we therefore plotted how mean score differed according to the six different demonstration types (Figure S.3), and according to the total number of rewarded model selections in a trial – 1, 2 or 3 (Figure 3.8). Breaking down the results in these additional ways did not reveal any patterns within the data which were not already captured in our other analyses. As in the analyses already reported, effects of age were apparent in the Temporal, Spatial and Temporal-Spatial conditions, particularly for higher scoring demonstrations.

Figure 3.8

Mean Child Score/3 by Age (Whole Years) and the Number of Rewarded Model Selections in a Trial for the Increased Information, Temporal, Spatial and Temporal-Spatial Conditions



Note. The dashed line at 1 depicts chance performance: the score expected if children were selecting chests at random and not using the social information. The dashed line at 2 depicts the highest single model score on any trial.

Discussion

Capitalising on increases in population size by utilising information from multiple models may be an important mechanism by which improved traits are introduced into human populations. We reasoned that exploiting this increase in information may require human-unique cognitive capacities because a larger cognitive load is likely generated when information from more than one source is separated in time and/or space. We thus presented children with Temporal, Spatial and Temporal-Spatial displacement conditions which emulated real-life constraints on information use and undoubtedly increased the cognitive load associated with obtaining maximum benefit from information provided by

multiple models. Our novel method allowed investigation of the ability to utilise this displaced information from two models, and moreover, children's potential to outperform the best available information within this multiple-model population. We predicted that this would be possible across a wider range of conditions as age, and thus cognitive capacity, increased. Such a result would suggest that enhanced cognitive abilities may better enable humans to exploit information provided by multiple models, despite the associated challenges, to generate new trait variants and drive cumulative culture.

Our results were broadly consistent with the above predictions. Firstly, in our least cognitively challenging control (Increased Information) condition, children utilised rewarded and unrewarded information at high levels right across our age range (5-10 years) and score did not differ according to age. Moreover, all ages evidenced outperformance of the highest single scoring model (i.e. the best available information in the population) and achieved a Level 3 on our continuum of potential for cumulative culture (p. 73). Children therefore displayed the ability to utilise information from the two models to the extent that they were able to generate an improved trait variant (a higher score) when the information was not displaced in time or space (i.e. when the information from both models remained present for use within the same search space, presumably creating only minimal cognitive load). In line with results by Subiaul et al. (2015), this demonstrated that even the youngest children did not find using information from two models difficult per se – rather, any difficulty in making the most advantageous use of information from multiple models may lie in the extra cognitive resources required to store, retrieve, and spatially translate information. In real-life social learning scenarios, we would expect the information from multiple models to be displaced in time and/or space therefore, although this result confirms that children can use information from two models in theory, it probably reveals little regarding children's propensity to do so in real-life.

Secondly, also in line with our predictions, we found that score and repeating of rewarded selections increased with age in the conditions which presented a greater cognitive load (Spatial, Temporal and Temporal-Spatial). The youngest children's performance was particularly weak in the Spatial and Temporal-Spatial conditions: 5-year-olds were the only age group with a lower score in both of these conditions compared to the control condition, and 6 and 7-year-olds scored more poorly in the Temporal-Spatial condition. However, older children's (8-10 years) score did not differ from the control in any condition. Additionally, older children (7-10 years) outperformed the highest scoring model and thus showed evidence of cumulative improvement across a wider range of conditions than younger children – in the Temporal, Spatial and Temporal-Spatial conditions (Level 3 on our continuum of potential for cumulative culture). Although 6-year-olds outperformed 5-

year-olds, their performance did not match that of 7-10-year-olds in the Temporal-Spatial condition – they repeated significantly fewer rewarded selections in comparison to the Increased Information condition and achieved Level 2. 6-year-olds therefore showed evidence consistent with the potential for cumulative culture in the Increased Information, Temporal and Spatial conditions only. However, 5-year-olds performed the most poorly – they appeared to gain little benefit from the presence of multiple models (achieving only Level 1) in either the Spatial or Temporal-Spatial conditions. They performed better in the Temporal condition (achieving Level 2) but still not to the extent that an improved trait variant was created.

As far as we are aware, our study is the first to demonstrate that children's ability to utilise information from multiple models may be affected by the types of cognitive constraints which exist in real-life social learning scenarios. Moreover, we are the first to show that children's ability to outperform the highest scoring single model in a multiple-model population, showing the potential for cumulative culture, changes with development (age 5-10) and the level of cognitive load (task constraints) presented. This is consistent with our previous work (Wilks et al., in press, Chapter 2) in which we found that children's capacity to use the social information provided by one model depended both on age and memory constraints (which differed across two task conditions). We thus postulate that, in any given experimental task or real-life scenario, the constraints on accessing information (e.g. temporal or Spatial separation), and cognitive ability, will determine whether these constraints can be overcome, allowing the information to be used and cumulative improvements to occur. We would expect that the ability to overcome particular constraints will differ depending on the cognitive abilities of the population e.g. adults, children at different stages of development, or nonhumans. In the case of accessing information from multiple models, this hypothesis is also supported by the adult experimental literature. Kempe and Mesoudi (2014) found that adults could integrate spatially separated information, leading to cumulative culture. Yet studies by Caldwell and Millen (2010) and Fay et al. (2019) highlighted that there may be constraints on adults' ability to integrate information if they do not have sufficient time in which to process it or the memory load is too great.

So why was the performance of 5-year-olds in our displacement conditions (and 6-year-olds, in the Temporal-Spatial condition only) poor in comparison to that of the other age groups? We postulate that the increased cognitive load presented by the temporal, spatial and temporal-spatial separation of the models' selections was difficult to overcome at this stage in a child's cognitive development. In our Temporal displacement conditions, age-related improvements to working memory (Best & Miller, 2010; Diamond, 2013; Garon et al., 2008) may have better enabled older children to hold the information provided by both

models in mind whilst deciding how best to utilise it. Furthermore, children aged 6 plus, with more developed metacognitive abilities and experience in formal schooling (Bryce & Whitebread, 2012), may have been better equipped to utilise metacognitive storage processes such as mnemonic devices (Jurowski et al., 2015) or inner speech (Carruthers, 2013; Cowan, 2008) to lessen cognitive load, and in particular the memory load. Such strategies have been described as “cognitive offloading” and children’s tendency to use (Armitage et al., 2020) and devise (Bulley et al., 2020) these increases with age. Children may have also shown increased metacognitive monitoring (e.g. judging task difficulty, self-questioning) and control (e.g. planning, changing strategy) with age (Bryce & Whitebread, 2012). Moreover, advancements in children’s metacognitive understanding of how perceptual access leads to knowledge at around age 6 may have resulted in increased task success from this age (although 6-year-olds did also struggle to make full use of the available information in the Temporal-Spatial condition with the highest cognitive load). Studies have shown that, until age 6, children have difficulty understanding that partial perceptual information leads to incomplete knowledge. As a result of this, prior to this age, children frequently overestimate their own knowledge (Kloo et al., 2017; Rohwer et al., 2012; Sodian & Wimmer, 1987) and that of others (Chandler & Helm, 1984; Taylor, 1988). When utilising information from two models, an appreciation of the consequences of partial exposure may have contributed to the higher performance of children aged 6 plus. For example, this insight may have encouraged children to be selective in their use of information from each model.

We may expect the performance of nonhumans to be similar to, or poorer than, that of 5-year-olds due to limitations in accessing and using the displaced information. Firstly, with respect to working memory, it is difficult to directly compare humans and nonhumans due to the small amount of comparative research available for analysis. However, it appears that, although some nonhuman mammals may have similar storage capacities, humans can represent concepts within memory differently (e.g. through the use of mnemonic devices, Jurowski et al., 2015, or inner speech, Carruthers, 2013; Cowan, 2008) and likely have a better ability to deploy attention and resist interference (for a systematic analysis of evidence to date see Carruthers, 2013). It is therefore likely that when presented with temporally displaced information nonhuman primates would perform similarly to young children, who are still developing such capacities, and would not be expected to produce an improved variant under temporal displacement.

Secondly, regarding metacognition, there is growing evidence from naturalistic information-seeking (Call & Carpenter, 2001) and uncertainty monitoring (Smith et al., 1997) paradigms, that some nonhumans (e.g. chimpanzees, orangutans, Bohn et al., 2017; Call &

Carpenter, 2001, gorillas, bonobos, Call, 2010, and rhesus monkeys, Couchman et al., 2010; Beran et al., 2006; Hampton et al., 2004) may have the ability to accurately monitor their own uncertainty and make simple responses according to this (Beran et al., 2012). However, these tasks require more limited information processing than that we have presented – the rewards are not split across multiple locations and the required responses are simpler. The ability to make such information-seeking and uncertainty monitoring responses is unlikely to allow for devising and using metacognitive strategies as may increase information use in this task, e.g. the aforementioned memory aids such as mnemonic devices or inner speech. Furthermore, the representational nature of this uncertainty is unknown (Beran et al., 2012; Carruthers, 2008, 2009) and (considering findings regarding animal mindreading, Call & Tomasello, 2008; Heyes, 2015) it seems unlikely that nonhumans explicitly represent “I know” (or a non-linguistic equivalent of such a belief) as humans do. This may limit their ability to make appropriately selective responses in more subtle situations (e.g. that presented in this study) in which partial information is available from multiple sources. Nevertheless, we might expect some nonhumans to be able to utilise information from multiple models under more limited circumstances, such as when the information is not displaced (e.g. in our Increased Information condition or that presented in Subiaul et al., 2015). This demands further research.

As noted in our introduction, making maximal use of multiple pieces of information separated in space likely requires one to translate the observed information to one’s own bodily frame of reference. This would be expected to increase the associated cognitive load due to requirements such as mental translation (Frick et al., 2013; Levine et al., 1999; Shepard & Metzler, 1971) and the ability to understand dual representation (DeLoache, 1989, 1991, 2000). In our Spatial and Temporal-Spatial conditions these skills are likely to be required for understanding that the models’ selections relate to each other and to comparative locations in one’s own search space. Moreover, we expect these will be needed for accurately translating relevant information (rewarded selections) from a model’s search space to one’s own. Research has shown that mentally moving visual information is cognitively taxing – e.g. there is a positive, linear relationship between the time taken to mentally rotate a shape and the angle through which it must be rotated (Shepard & Metzler, 1971). Such ability is considered to be present from age 5 (Frick et al., 2013; Iachini et al., 2019; Marmor, 1975; evidence in younger children is inconsistent, Estes, 1998; Frick et al., 2009, 2013) and to continue to develop throughout childhood and into adolescence (Kail et al., 1980). This developmental trajectory thus supports our finding that utilising the spatially separated information was challenging for the youngest children, especially when combined with the added constraint of temporal separation. It is likely that nonhuman primates would

perform similarly due to the increased cognitive load required to translate information from multiple locations separated in space. However, further research would be needed to ascertain this as there is some evidence that nonhumans can engage in mental rotation (Köhler et al., 2005; Stich et al., 2003; Vauclair et al., 1993), although currently no evidence that this is homologous to the process in humans (Carruthers, 2013).

We have shown that children aged 6-10-years-old can utilise information from multiple models to generate improved variants under ecologically valid constraints. However, it is unlikely that all new cultural traits are derived by combining information across different models, and some traits may not lend themselves to such combinatory mechanisms. To reiterate, we do not argue that the kind of combinatory social learning investigated in this study is the only method by which improved variants may arise within populations, but rather, that human cognition may enable our species to exploit information from multiple models and thus engage in this method.

Theories which posit that population size underlies human cumulative culture (Henrich, 2004; Powell et al., 2009; Shennan, 2001) have been set up in opposition to cognitive explanations. However, we have shown that deriving full benefit from the Increased Information provided by multiple models may be dependent on the ability to store, manipulate and retrieve this information. It therefore follows that the effects of population size on cumulative culture may be inherently dependent on the cognitive abilities of the learners. Population size and cognitive explanations should not be viewed as being in opposition with one another. Rather, they are intrinsically linked.

Chapter 4: Who Knows? Does Children's Developing Understanding of Others' Minds Affect Their Social Learning Strategies? Part 1.

Introduction

Social learning strategies (SLSs) are heuristics describing the biased use of information derived from others. Originally predicted by theoretical models (Laland, 2004), there is now empirical evidence of SLSs in humans (adults and children) as well as an array of nonhuman species. On a rudimentary level, SLSs can be thought of as names for “strategies” scientists have observed when individuals engage in social learning, and these fall into three categories: what information should be copied, when one should copy, and who should be copied (Heyes, 2016; Kendal et al., 2018; Morgan et al., 2012). This selective copying can aid retention of beneficial traits within populations, thereby increasing fitness. However, as SLSs are present in both animals and humans it seems that they alone cannot account for the selection and retention of beneficial traits to the extent that the functionality of the trait is continuously increased over generations of social transmission. This process, termed cumulative culture or the *ratchet effect*, seems to be largely restricted to humans, therefore if SLSs are at all responsible we might expect SLS that enhance culture in this manner to be somewhat different to SLSs exhibited by nonhumans.

Despite ongoing research interest in SLSs, and evidence of their widespread use, the cognitive mechanisms underlying their deployment in both humans and nonhumans are unconfirmed. The word “strategy” implies awareness and deliberate use, but it would be incorrect to assume individuals are aware that they are using SLSs. In fact, it has been proposed that many SLSs, in non-humans and humans alike, rely on associative learning mechanisms and so are applied heuristically (Heyes, 2016; Kendal et al., 2018). Such SLSs have been described as “planetary” because they would not be implemented with deliberate intent and would be strategies only in the minds of their observers, much like the laws governing planetary motion (Heyes, 2016). An example of such a strategy, and one which we utilise in the study presented here, involves copying the behaviour of most individuals – “copy-the-majority”. Although this strategy is widely documented in humans (both adults, Asch, 1951, and children, Burdett et al., 2016; Haun & Tomasello, 2011; Van Leeuwen et al., 2018; Wilks et al., 2015), and may or may not be applied with awareness in these cases, it has also been demonstrated in nonhumans (Haun et al., 2012; Pike & Laland, 2010), cases which Heyes would certainly attribute to associative learning. It is easy to see how multiple demonstrators could offer a stronger associative cue, compared with a single demonstrator, regardless of the learner’s level of insight into their own use of social information.

Conversely, humans may be the only species with a mental representation of rules governing “who” and “when” we should copy; these SLSs could be considered “cook-like” because they would be reportable strategies one would be explicitly aware of employing deliberately, as a cook is aware of the rules they adhere to when cooking. Such “special” social learning strategies, termed explicitly metacognitive SLSs, have been proposed to require higher-level (System 2) processing, hence their anticipated confinement to humans (Heyes, 2016). Similarly, Wood et al. (2013) have proposed that higher-level cognitive abilities, such as perspective taking and an understanding of higher-order mental states, may enable humans to use SLSs flexibly, selecting the most appropriate SLSs for use in particular contexts and aiding selection of the most suitable models. More recently, it has been hypothesised that humans may be able to use social information more selectively than other species based on explicit metacognitive (by this definition encompassing both introspection and mentalising) understanding, which is unique to humans. SLSs utilising such metacognitive abilities, which we will refer to as understanding-based, explicitly metacognitive SLSs, would differ slightly from the metacognitive SLSs proposed by Heyes (2016, 2018) because, in addition to an agent knowing that they are using a SLS, the SLS itself would result from explicit metacognitive understanding of the current available information i.e. the agent would devise the strategy based on their explicit metacognitive understanding, and understand why they were using it. For example, an ability to explicitly represent understanding of knowledge and ignorance in others (a component of theory of mind) could allow humans to recognise a more knowledgeable individual and selectively employ an understanding-based, explicitly metacognitive SLS to copy information from that individual. This understanding that one individual has specific, task-relevant and up-to-date knowledge could enable more general-purpose, default SLSs to be over-ridden and information from this model to be copied in preference. Additionally, explicitly understanding one’s own knowledge/ignorance, and what kind of evidence is needed to gain any required knowledge, may aid in understanding that individual knowledge is insufficient and therefore another individual with superior knowledge should be identified and copied. Utilising human-unique explicitly metacognitive processes to use social information in this flexible manner (i.e. to copy the individuals from which the most benefit can be derived) has been argued to facilitate more optimal information gain than reliance on general-purpose SLSs, which are not fine-tuned to the specific situation. Explicitly metacognitive processes may therefore enable cumulative culture through allowing ratchet-like advances as those with the best, most up-to-date information are selectively, and precisely, copied (the explicitly metacognitive cumulative culture hypothesis, Dunstone & Caldwell, 2018). This hypothesis has yet to be tested empirically.

If human-unique social cognitive/metacognitive capacities are essential for the kind of selective social information use we have described above, we would expect there to be an association between milestones in human development of these capacities and ability to use social information selectively. Thus, a shift from use of the proposed nonhuman-like (associative) social learning strategies to more adult-like (understanding-based, explicitly metacognitive) social learning strategies at a particular point in a child's cognitive development would support the hypothesis that human social learning strategies are different to those of nonhumans and that this may be due to human-unique cognitive capacities (Dunstone & Caldwell, 2018; Heyes, 2016; Wood et al., 2013). In the next section we will therefore briefly discuss milestones in children's understanding of the mind which we think may contribute to such a shift in how social information is used. In our study we aimed to utilise these milestones to test the theory that some human social learning strategies may be fundamentally different to those of nonhumans.

The human capacity to understand the minds of others has been a subject of intense study since Premack and Woodruff (1978) set out to determine if the chimpanzee "imputes mental states to others". This subject is still under debate; however, current research suggests that the chimpanzee understanding of others' minds is not more advanced than that of a human infant. In a non-verbal analogue of the gold-standard false belief test (Wimmer & Perner, 1983), Call and Tomasello (1999) found that nonhuman apes did not appear to attribute false beliefs, whereas young children's performance on this test was equivalent to their performance on the standard, verbal false belief test (Wimmer & Perner, 1983). Further studies have confirmed that great apes are unable to pass false belief tests requiring explicit behavioural responses (Hare et al., 2001; Kaminski et al., 2008; Krachun et al., 2009). However, a recent anticipatory looking, false-belief study (Krupenye et al., 2016) has indicated that great apes, like human infants in similar studies (Clements & Perner, 1994; Southgate et al., 2007), may have an implicit understanding of false beliefs.

In humans, the false belief test (Wimmer & Perner, 1983) has been used extensively to probe explicit theory of mind, more specifically – understanding of false belief in others. The preoccupation with this task is likely because passing it seems to represent a significant developmental milestone in children's understanding of the mind: the transition from failure (generally age 3 and below) to success (at around age 4) occurs over a very narrow time period (Doherty, 2009). Passing the test suggests a rather sophisticated understanding of others' beliefs and the basis on which these beliefs are formed – an individual holds a belief which is justified by their experience yet the belief is distinct from "knowledge" because it differs from the current, true reality known to the child. An understanding of false belief may therefore facilitate more selective social information use due to greater appreciation of when

the belief another individual has is true and based on evidence. Of course, this is difficult to investigate – we cannot remove an individual's ability to understand false belief – but if we found a shift in the way social information is used around the time children are able to pass the false belief test this would indicate that this cognitive capacity may be a contributing factor.

Whilst passing the false belief test appears to represent a particularly significant milestone, there is evidence that children's understanding of others' minds is becoming increasingly sophisticated prior to age four. Research under the "theory of mind" umbrella has probed children's understanding of how evidence is related to knowledge in others (and also in themselves – termed metacognition). Most commonly, this concerns how visual perceptual access to the contents of a container (i.e. looking inside) leads to knowledge of those contents. The results of these studies vary according to who has perceptual access, the child themselves or another individual, and the kind of task used. Most relevant to the current discussion is research showing that children begin to understand the link between another individual's (child, adult or puppet) visual access and the knowledge that individual then gains from age 3 or 4 (Pillow, 1989; Pratt & Bryant, 1990; Schmidt & Pyers, 2011; Wimmer et al., 1988). That is, children can answer questions of the type: "Does (individual's name) know what is inside the box?" correctly (Pillow, 1989, Experiment 1; Pratt & Bryant, 1990, Experiments 2 and 3; Wimmer et al., 1988) or can actively choose which of two individuals, one with and one without visual access, knows the contents of a box (Pratt & Bryant, 1990, Experiment 1; Schmidt & Pyers, 2011) or should be asked what colour the object hidden in the box is (Pillow, 1989, Experiment 2). It has also been claimed that children of this age can make a search decision, and find a hidden reward, based on their understanding that an individual who has visual access should be copied in preference to an individual who does not (Palmquist et al., 2012; Povinelli & deBlois, 1992). However, the training involved in these tasks does not rule out solving via associative processes.

The finding that children are beginning to understand the relationship between another individual's visual access and this individual's knowledge state holds true regardless of a child's own knowledge state. This has been examined in the following task types: total exposure (the participant child has had visual access to the contents of the container themselves), total ignorance (the participant child has not had visual access to the container's contents), and partial exposure (the participant child has been exposed to the potential contents of the container). In partial exposure tasks the observer knows (for example) that either a red or a blue ball has been transferred to a container, but they have not had visual access to the contents of the container post transfer (Pillow, 1989; Schmidt & Pyers, 2011). It therefore appears that children are beginning to appreciate the relationship

between an individual's visual access and their knowledge state prior to the age at which they understand that an individual may hold a false belief. Furthermore, there is some evidence that children of this age can utilise their understanding of seeing and knowing in others to help them make simple decisions (Palmquist et al., 2012; Pillow, 1989, Experiment 2; Povinelli & deBlois, 1992; Pratt & Bryant, 1990, Experiment 1; Schmidt & Pyers, 2011). Yet, the extent to which this understanding enables selective use of social information has not been fully investigated. Our study will build upon work which has investigated how a child's understanding of what another individual sees and knows affects their decision making. We investigate whether any change in the way social information is used coincides with the age at which children's understanding of seeing and knowing increases.

Although children aged 3 are developing an awareness of the association between an individual's perceptual access and knowledge, their understanding of this concept appears to be limited. That is, children of this age may associate seeing with knowing without understanding the information which can, or cannot, be gained from different types of perceptual access (O'Neill et al., 1992; Povinelli & deBlois, 1992). In fact, it is not until age 5 or 6 that children begin to pass tests requiring an understanding that partial information is often insufficient for an individual to gain all the knowledge required. In the "Doodle task" (Chandler & Helm, 1984; Taylor, 1988) children are shown a large picture followed by a restricted view image/images of the same picture. For example, the full image might be a drawing of an elephant, and the smaller segment only the lines forming a single leg. They are then asked if a puppet who has only seen the restricted view image/s will know what is in the full picture. Children do not recognise that the puppet does not have enough information to know what the full picture depicts until age 5-6. Moreover, children younger than this commonly over-estimate the knowledge the puppet has from its limited visual access. In a similar vein, research has shown that until age 4/5 children do not seem to understand (for both themselves and others) that different sensory experiences enable the establishment of different types of knowledge. For example, O'Neill et al. (1992) presented children with object pairs which were identical in all but one dimension and could therefore be differentiated using one of the senses only e.g. sight to differentiate two objects which differed only in colour, and touch to differentiate two objects which differed only in weight or some other tactile property. On each of several trials, one object from one of the pairs was hidden and two puppets, one restricted to visual access and one to tactile access, attempted to identify properties of the object. Children were asked if the puppets' actions (looking/feeling for visual/tactile access respectively) meant that the puppets were aware of the properties of the objects which allowed differentiation (the visual or tactile properties). Only by age 4 and 5 were children able to accurately access when the puppets would have

knowledge of the differentiating properties. As in the Doodle task, (Chandler & Helm, 1984; Taylor, 1988), 3-year-olds commonly over-estimated the knowledge which could be gained from looking i.e. they often stated that the puppet with visual access had knowledge of the properties of the tactile objects. As for the social cognitive developments previously discussed, our study should shed light on whether the capacity to understand how evidence relates to knowledge, to the extent outlined above, contributes to developmental changes in the way in which social information is used.

We have highlighted developmental milestones in children's understanding of the mind which may confer an advantage for selective social information use as posited in the explicitly metacognitive cumulative culture hypothesis (Dunstone & Caldwell, 2018). As previously stated, such developments may underlie any observed shift from nonhuman-like (associative) to adult-like (understanding-based, explicitly metacognitive) social learning strategies during human development. However, to our knowledge, no study has directly investigated whether such a shift in the use of social information occurs as understanding of the mind goes from the level of a human infant to that of an adult, or at what age this might occur (of course, this age may vary depending on the information itself and how it is presented – see Chapters 2 and 3). Identifying a shift in social information use during human development would support the hypothesis that humans use social information differently due to underlying cognitive mechanisms (Dunstone & Caldwell, 2018; Heyes, 2016; Wood et al., 2013).

We provided children, aged 3 to 8, and adults with social information as part of a game in which three uninformed demonstrators, and one informed demonstrator (animal puppets), searched two locations for hidden objects. Children needed try and determine the objects' location but did not observe the outcome of the demonstrators' search attempts until the end of the game, so individual learning about the correct location was not possible. The task was devised such that use of a strategy based on understanding of others' minds (an understanding-based, explicitly metacognitive strategy) would enable task success. The informed demonstrator had visual perceptual access to task-relevant knowledge in that they could see the contents of both hiding locations. It was thus possible for children to use their understanding of this individual's mind, specifically that they had knowledge due to visual access, to locate the hidden objects. We pitted this strategy against one which we expected to dominate in the absence of this understanding of the informed demonstrator's mind and thus a strategy which would not enable task success. For this purpose we utilised the common, adaptive SLS to "copy-the-majority" which has been identified in humans (adults, Asch, 1951, and children, Burdett et al., 2016; Haun & Tomasello, 2011; Van Leeuwen et al., 2018; Wilks et al., 2015) and a range of nonhuman species (Haun et al., 2012; Pike &

Laland, 2010). This strategy can easily be applied heuristically without an understanding as to why it is effective therefore if children did not understand that the informed individual had additional knowledge it was expected that they would resort to use of this default strategy and copy one of the three uninformed demonstrators. We aimed to examine whether there was an age-related shift from use of this majority bias strategy to a strategy based on the attribution of knowledge. Finding such a shift in social information use might indicate when in development children begin to explicitly consider others' knowledge in directing their social learning.

The goal of our task, both for the human participants and the informed/uninformed puppet/s, was to find objects hidden inside plastic eggs in one of two locations. Three individuals (the uninformed majority without visual access to the objects hidden inside the eggs) were observed choosing the same, incorrect location. However, the one informed individual was observed gaining additional task-relevant perceptual information (opening the eggs) upon each of three search attempts/choices and could therefore be considered to know in which location the objects were hidden following their first search attempt. In Condition 1 this first search attempt was the incorrect location and the following two attempts the correct location, but in Condition 2 all three search attempts were directed towards the correct location. We included these two conditions because we did not know which of the informed demonstrator's actions would be more salient i.e. we thought it possible that observing the informed demonstrator switch location in Condition 1 may provide an enhanced cue as to this demonstrator's knowledge compared to opening the eggs alone (as happened in both conditions). Moreover, as our task was the first of this type to investigate whether there is a switch in SLS use during development, it was possible that participants might not interpret the task as intended. For example, we did not know for certain that participants who did have an explicit understanding of the informed demonstrator's knowledge state would actually choose to copy the informed demonstrator. Adult participants were used to test this assumption. After watching the informed/uninformed individual/s, participants were required to choose one of the two locations themselves, i.e. the one which they believed contained the hidden target objects. They were not made aware of whether the informed/uninformed individual/s had chosen the location containing the target objects until the end of the experiment because we did not want participants to base any strategy on visible success/failure of the models, but rather on their interpretation of the models' actions.

We predicted that we would observe a shift from more nonhuman-like (associative) task responses (i.e. copying the majority of uninformed individuals) to more adult-like (understanding-based) responses (i.e. copying the informed individual) as children's explicit metacognitive abilities increased with age. We had no a priori expectations regarding an

exact developmental stage at which this shift might occur because we do not know which explicit metacognitive developments might be most critical e.g. understanding of false belief, how evidence relates to knowledge, partial knowledge exposure, or possibly other related concepts.

Method

Participants

We collected data from 192 children aged 3 to 8. Children were recruited from two nurseries, three primary schools and Blair Drummond Safari Park (Stirling, Scotland, UK). In schools we worked with the year groups Primary 1 – Primary 3 to cover our required age range. Note that it was not uncommon for children of the same age to be in different school year groups due to the flexibility in school starting age in Scotland. 19 children were excluded for the following reasons: misunderstanding of task instructions (n=2, aged 3 and 6, female), lack of response to questions (n=2, aged 3, male and female), lack of concentration on the task (n=2; aged 4, female; and 5, male), lack of interest in the task (n=2; aged 4, male; and 8, female), recognised developmental delay (n=1, aged 6, male), experimenter error (n=3; aged 3, male; 6, male; and 8, female), the child had previously watched other children perform the task (n=2, aged 7, female), and selection to participate (by the school) conducted in a biased manner (n=5; aged 8; two male, and three female). The final sample consisted of 173 children aged 3 to 8 (M= 71 months, range= 40 – 108, SD= 20, 84 male and 89 female). 87 were assigned to Condition 1 (45 male and 42 female), and 86 to Condition 2 (39 male and 47 female). There were between 13 and 17 children of each age group in each condition (Table 4.1).

Table 4.1

Total Number of Participants (Aged 3-8 Years and Adult) and Number Who Chose the Correct/Incorrect Basket in Conditions 1 (Adult and Child Data) and 2 (Child Data Only); % in Brackets

Age (whole years)	Total number participants		Number participants who chose correct/incorrect basket (% of total participants in brackets)			
	Condition 1	Condition 2	Condition 1		Condition 2	
			Correct	Incorrect	Correct	Incorrect
3	13	13	9 (69.2)	4 (30.8)	7 (53.8)	6 (46.2)
4	16	17	9 (56.3)	7 (43.7)	7 (41.2)	10 (58.8)
5	15	14	7 (46.7)	8 (53.3)	8 (57.1)	6 (42.9)
6	13	14	5 (38.5)	8 (61.5)	8 (57.1)	6 (42.9)
7	16	14	8 (50.0)	8 (50.0)	9 (64.3)	5 (35.7)
8	14	14	11 (78.6)	3 (21.4)	7 (50.0)	7 (50.0)
adult	20	NA	12 (60.0)	8 (40.0)	NA	NA

24 adults were recruited from the University of Stirling via advertisement on the University intranet. Four adults were excluded for the following reasons: misunderstanding of task instructions (n=1, male), lack of concentration on the task (n=1, male), and experimenter error (n=2, female). The final sample therefore consisted of 20 adults aged 19 to 55 (M= 380 months, range = 230 – 660, SD= 117, 7 male and 13 female). All adults were assigned to Condition 1.

Ethics statement: This research was approved by the University of Stirling, General University Ethics Panel (references: GUEP 374 and 374A). Written, informed consent was obtained from all adults and the parent or guardian of all children prior to their participation. Children were asked if they would like to participate, were continuously monitored for assent and were rewarded with a sticker regardless of task completion. Adults were provided with monetary compensation for their time and were informed that they could stop participating at any point during the experiment.

Materials

Large farmyard animal puppets – a cow, dog, duck, hen and pig (The Puppet Company©) were used as demonstrators in social information trials (see Figure 4.1). The hen always acted in a fixed role, and the role of the four remaining puppets was randomised. A large, soft toy fox (Wild Republic©) was also used in a fixed role. The choice stimuli were two wicker baskets (diameter base: 125mm, diameter opening: 175mm, height: 175mm) half filled with tissue paper, each containing 14 gold, plastic eggs (diameter centre: 40mm, height: 60mm). One basket contained eggs with toy chicks (height: 30mm, width: 15mm, length: 15mm) inside, and the other basket contained eggs with imitation yolks inside which were made from balls of yellow tissue paper of the same size and weight as the chicks. The baskets were placed onto a laminated board (length: 412mm, width: 392mm) depicting a farmyard scene. Six gold, plastic coins (diameter: 23 mm, height: 5mm) were used by the puppets as imaginary payment, and a further six were used by the participant and presented to them inside a small, wooden chest (length: 82mm, width: 53mm, height: 46mm).

Figure 4.1

Experimental Setup Including the Following: Hen Puppet, Two Wicker Baskets, Plastic Eggs, Soft Toy Fox, Laminated Board Depicting a Farmyard Scene and One of The Four Informed/Uninformed Farmyard Animal Puppets (the Cow)

**Procedure**

At the safari park child testing was carried out in a public space, separated from other visitors by a table and a gazebo. At schools/nurseries testing took place in a quiet area of the classroom, or in a separate room. At the safari park only, less confident children were accompanied in the testing area by a parent or guardian, who was instructed not to provide the child with any task-relevant assistance. In both locations the experimental task was carried out on a table-top (see Figure 4.1) with the child seated opposite the experimenter

(operating the puppets) and next to the research assistant (live coding the child's responses). A verbal script was used by both the experimenter and the research assistant (see Appendix "Experimenter Verbal Scripts"). Prior to starting the experiment, the laminated board was placed on the table in front of the child's seat, and the two wicker baskets (each containing 14 eggs, one basket those with chicks inside and the other those with yolks) positioned towards the back of the board. The fox was placed in between the baskets to create the impression that it was guarding them (Figure 4.1). The testing procedure for adults was identical to that for children except that adults were tested individually, in a small room at the University of Stirling; it was explained to them that the task was designed to be engaging for young children. Adults were assigned to Condition 1 only (further detail below).

All participants were introduced to the task through a short interaction with the hen puppet and fox. Child participants were asked if they would like to play a game (see Appendix "Experimenter Verbal Scripts" for exact language used throughout) and adults were asked whether they would like to take part in a task we were running with children. All participants were informed that the goal was to help the puppet named Hen find her lost chicks. Participants were told that the hen had two types of eggs – some with chicks inside and some with yolks inside ("for Mr Fox to eat"). They were shown examples of these eggs, which they were instructed to open, and told that the hen had placed all her eggs with chicks inside into one basket, and those with yolks inside into a second basket. Participants were informed that the fox had taken both baskets and swapped them around so that the hen no longer knew the location of her chicks and therefore needed help to find out which basket they were in. Additionally, participants were told that if any baskets were to be looked inside then the fox demanded payment in the form of coins, and that some "animal friends" (the cow, dog, duck and pig puppets) were going to help find out which basket contained the chicks. A social information trial followed by a test trial was then conducted.

Social Information Trial. A participant observed three animal puppets (from the cow, dog, duck and pig) each pay the fox one coin and remove an egg from one of the two baskets in their search for the chicks. These three demonstrators were considered to be uninformed because they were each selecting a basket without having visual access to the contents of any of the eggs inside it. Each of these demonstrators selected the incorrect basket (i.e. that which did not contain the chicks) and the basket was therefore selected on three consecutive occasions. Upon each selection the uninformed demonstrator would take an egg and place it onto the laminated board without opening it. Either before or after the uninformed demonstrators, a participant would observe the one remaining, informed puppet (introduced as having three coins) remove an egg from one of the baskets and pay the fox

one coin before proceeding to peek into it without allowing the child to see the contents. The closed egg was placed onto the laminated board and the puppet would then be observed making two further selections, each time paying the fox one coin and visibly peeking into the removed egg before placing it onto the board. The three locations selected by this puppet depended on the experimental condition the adult/child had been assigned to, and were as follows: Condition 1 – first selection the same (incorrect) location as that selected/to be selected by the uninformed demonstrators, and the second and third selections the alternative (correct) location i.e. that containing the chicks; Condition 2 – all three selections the alternative (correct) location. We considered this demonstrator the “informed” demonstrator because it was the only puppet to peek inside any of the eggs, and therefore the only puppet with visual access to their contents.

The condition to which children were assigned was determined according to a combination of factors, including the maintenance of balanced numbers across conditions within particular age groups, and ease of management by the experimenter (only recruiting to a single condition during particularly intense testing sessions where multiple participants were to be tested consecutively). All adults were assigned to condition 1. Prior to testing, and for each participant number, we randomly assigned the following: which of the four animal puppets would take the role of the informed demonstrator, which of the two baskets would be selected by the uninformed demonstrators and would therefore contain the yolks (which thus determined which basket/s would be selected by the informed demonstrator), and which of the three uninformed demonstrators would be asked about in the memory question (see the below “test trial” section below for further details). We counterbalanced whether the informed demonstrator, or uninformed demonstrators, would make their selections first by alternately assigning each two participant numbers to “informed first” or “uninformed first”. We then counterbalanced whether the informed, or randomly selected uninformed, demonstrator was asked about first in the memory questions by alternately assigning each participant number to “ask about informed first” or “ask about uninformed first”.

Test Trial. Following the social information trials, the research assistant introduced the participant to a small chest containing six gold, plastic coins. The participant was informed that the fox would accept these coins as payment for a whole basket of eggs, and told to select the basket they believed the chicks were inside in order to rescue all of them for the hen – “Which basket should we buy to rescue the chicks?” Both verbal and non-verbal responses (i.e. pointing at one of the baskets) were accepted, and the participant was

told that they would be able to look inside their chosen basket shortly. The assistant then asked: “Why did you choose that basket?”

Participants were asked a further three questions to assess their memory of the social information trials. These questions were as follows: “Which basket did [informed demonstrator puppet name] look inside on their last turn?” “Which basket did [uninformed demonstrator puppet name] look inside?” and “Which animal had the most coins?” As stated above, we counterbalanced the order in which the first two questions were asked and randomised which of the three uninformed demonstrators was asked about. Whilst these memory questions were being asked, the experimenter held up the relevant puppet/s to aid memory and to enable participants to respond to the last question via pointing (if desired). Pointing at one of the two baskets was accepted as an answer to the first two questions.

Lastly, the participant was invited to look inside their selected basket in order to discover whether they had found the chicks. For children only, those who had chosen the wrong basket (i.e. eggs with yolks inside) were then invited to ask the fox to swap this basket so that they had the satisfaction of helping the hen. Total testing duration was 10-15 minutes per child and 5-10 minutes per adult. The responses to all questions (verbal, or none verbal where appropriate) were live coded by the research assistant.

Results

All generalised linear mixed effects models (GLMM) were carried out with the logit link (binomial data, family = binomial) and the lme4 package (lmer function) using R (R Core Team, 2018). p -values < .05 were taken as statistically significant across all analyses (unless otherwise indicated). We endeavoured to keep random effects structures “maximal” where possible, following Barr et al. (2013). Where the maximal model resulted in non-convergent or singular fit models, random slopes were removed followed by random intercepts where necessary until a convergent, non-singular model was obtained.

Basket Choices

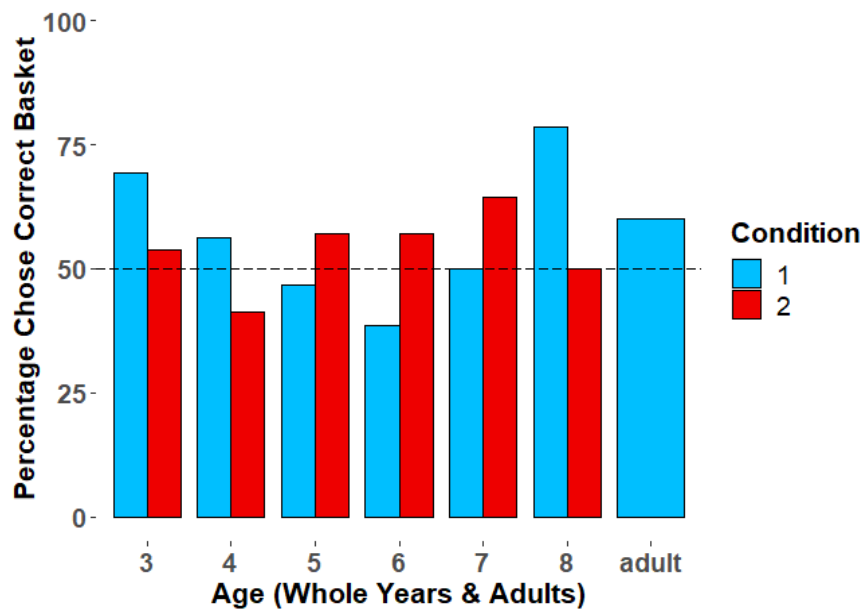
We were interested in whether there was a shift towards more adult-like, understanding-based task responses (i.e. copying the informed demonstrator’s correct choice rather than the incorrect choice of multiple uninformed demonstrators) with increasing age (3-8 years), and whether this differed according to task Condition (1 or 2). We therefore scored each child and adult participant as having made the correct choice (1), or the incorrect choice (0) and analysed the effect of child age and condition on “choice correct”. We thus performed a GLMM with choice correct (0 or 1) as the dependent variable and age (in days), Condition (1 or 2), and the interaction of age and condition as fixed effects. We also included a random intercept for participant number but no random slopes. Age was

centred and scaled to measure thousands of days, and condition was sum coded (condition 1 set to 1; Condition 2 to -1). There were no main effects of age ($b = 0.058$, $SE = 0.252$, $Z = 0.231$, $p = .817$), condition ($b = 0.057$, $SE = 0.153$, $Z = 0.373$, $p = .709$) or the interaction of age and condition ($b = -0.001$, $SE = 0.252$, $Z = -0.005$, $p = .996$) on choice correct. A chi-square test for independence was also performed on the above data but separating this into Condition 1 and 2 and using age in years. This confirmed that the percentage of participants choosing the correct/incorrect baskets (Table 4.1) was independent of age category in both Condition 1 ($\chi^2(5, N = 87) = 6.212$, $p = .286$) and 2 ($\chi^2(5, N = 86) = 1.911$, $p = .861$).

We did not find a shift towards choosing the same basket as the informed demonstrator with increasing age, as we would have expected if developments in explicitly metacognitive understanding enabled social learning strategies based on these developments to be devised. We therefore wanted to further explore how children's responses were distributed between the two baskets and compare these results to those from our adult sample i.e. were children choosing the same basket as the informed demonstrator above chance level at any age group? Or were children perhaps responding as if they were choosing a basket at random (i.e. neither of the baskets chosen above chance level)? We thus calculated the number and percentage of participants who chose the correct basket for each age and condition (Table 4.1 and Figure 4.2). Two-sided, exact binomial tests (Table 4.2) showed that the proportion of children who chose the correct basket was not significantly different to the chance proportion of 0.5 at any age group or condition (ages 3-8, all $p > .05$), supporting the hypothesis that children may have chosen a basket as if at random. However, at age 8, in Condition 1 only, the result of the binomial test was approaching significance ($p = .057$). Yet, somewhat surprisingly, adult participants failed to choose the correct basket significantly above chance levels (two-sided, exact binomial test $p = .503$). Adults therefore performed more poorly than expected – a point we will return to in our discussion.

Figure 4.2

Percentage of Participants Who Chose the Correct Basket by Condition and Age (Whole Years & Adults)



Note. The dashed line shows the percentage performance we would expect if participants were selecting a basket at random. All adults were assigned to Condition 1.

Table 4.2

Summary of Two-Sided Exact Binomial Tests for Conditions 1 (Adult and Child Data) and 2 (Child Data Only), Both for the Full Data Set and a Subset Which Included Only Participants Who Answered all Three Memory Questions Correctly

Condition	Age (whole years)	All participants		Participants (all memory questions correct)	
		Proportion chose correct basket	p value – binomial test (two-sided)	Proportion chose correct basket	p value – binomial test (two-sided)
1	3	.692	.267	1.000	.125
1	4	.563	.804	.500	1.000
1	5	.467	1.000	.375	.727
1	6	.385	.581	.250	.289
1	7	.500	1.000	.500	1.000
1	8	.786	.057	.889	.039
1	adult	.600	.503	.692	.267
2	3	.538	1.000	.333	1.000
2	4	.412	.629	.400	.754
2	5	.571	.791	.636	.549
2	6	.571	.791	.700	.344
2	7	.643	.424	.636	.549
2	8	.500	1.000	.545	1.000

Participants may have had difficulty remembering from which of the two baskets each of the four demonstrators removed an egg and whether each demonstrator was informed (peeked inside their retrieved egg), or uninformed (did not peek inside their retrieved egg). As already stated, we determined which participants were able to remember this information through three memory questions: “Which basket did [informed demonstrator puppet name] look inside on their last turn?”, “Which basket did [uninformed demonstrator puppet name] look inside?” and “Which animal had the most coins?”. We were therefore able to perform further analysis of the participants’ choices, as above, but on a subset of the data (Table 4.3) which included only those who passed all three memory questions. A significantly higher proportion of correct basket responses for these subsets, and/or an age effect not present in the full dataset, would indicate that difficulty with continuously monitoring/remembering the demonstrator’s selections may have impeded ability to make the correct choice.

Table 4.3

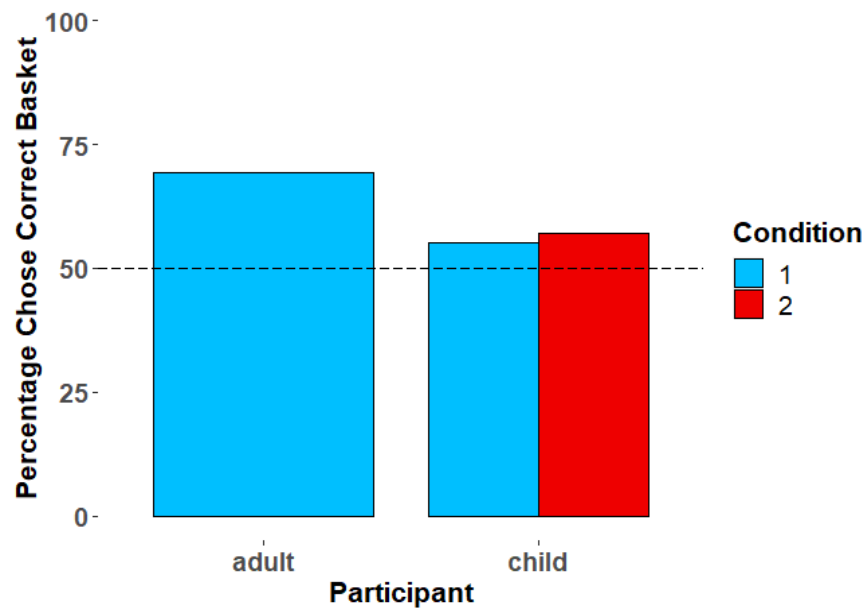
Number of Participants (Ages 3-8 and Adult) Who Got All Three Memory Questions Correct and Number of Participants Who Chose the Correct Basket Out of the Number of Participants Who Got all Three Memory Questions Correct for Conditions 1 (Adult and Child Data) and 2 (Child Data Only); % in Brackets

Age (whole years)	Number participants all memory questions correct/total participants (% in brackets)		Number participants who chose the correct basket/number participants all memory questions correct (% in brackets)	
	Condition 1	Condition 2	Condition 1	Condition 2
3	4/13 (30.8)	3/13 (23.1)	4/4 (100.0)	1/3 (33.3)
4	8/16 (50.0)	10/17 (58.8)	4/8 (50.0)	4/10 (40.0)
5	8/15 (53.3)	11/14 (78.6)	3/8 (37.5)	7/11 (63.6)
6	8/13 (61.5)	10/14 (71.4)	2/8 (25.0)	7/10 (70.0)
7	12/16 (75.0)	11/14 (78.6)	6/12 (50.0)	7/11 (63.6)
8	9/14 (64.3)	11/14 (78.6)	8/9 (88.9)	6/11 (54.5)
adult	13/20 (65.0)	NA	9/13 (69.2)	NA

In order to examine whether there was an age effect we performed a second GLMM, using the same dependent variable/fixed effects and random effect as that performed on the full data set. As in the previous model, there was no main effect of age ($b = 0.205$, $SE = 0.348$, $Z = 0.590$, $p = .555$), condition ($b = -0.041$, $SE = 0.198$, $Z = -0.207$, $p = .836$) or the interaction of age and condition ($b = -0.024$, $SE = 0.347$, $Z = -0.068$, $p = .946$) on choice correct. We performed further exploratory analyses including a Fisher's exact test which showed that the percentage of participants choosing the correct/incorrect baskets (Figure 4.3) differed according to age category in Condition 1 ($p = .040$), but not Condition 2 ($p = .737$). Further post hoc, Fisher's exact tests on the Condition 1 data (Table 4.4) revealed that the percentage of children aged 8 choosing the correct/incorrect basket differed from that at all other age groups ($p = .031$); children aged 8 only were more likely to choose the correct basket (89% aged 8 and 48% for all other age groups combined). However, when the Bonferroni correction for multiple comparisons was applied this result for the 8-year-olds ($p = .031$) was no longer significant at the new alpha level of .008. We also performed two-sided, exact binomial tests (Table 4.2), as for the full dataset, in order to determine if the proportion of children choosing the correct basket was significantly different to 0.5 at any age group. The proportion of children who chose the correct basket was significantly different to this chance level at age 8, Condition 1 only ($p = .039$ – see Table 4.2 for p values ages 3-7 [all $p > .05$]). As in the analysis of the full data set, adult performance was not significantly different to chance level ($p = .267$) and was therefore poorer than expected (Figure 4.3).

Figure 4.3

Percentage of Child and Adult Participants Who Chose the Correct Basket by Condition for a Subset of the Data Which Included Only Participants Who Answered All Three Memory Questions Correctly



Note. The dashed line shows the percentage performance we would expect if participants were selecting a basket at random. All adults were assigned to Condition 1.

Table 4.4

Summary of Post Hoc, Fisher's Exact Tests on the Percentage of Children Who Chose the Correct/Incorrect basket (see Table 4.2) for Each Age Group Against All Other Age Groups Combined (Condition 1, Subsetted Data Including Only Participants Who Answered all Three Memory Questions Correctly)

Age (whole years)	<i>p</i> value – post hoc, Fisher's exact tests
3	.117
4	1.000
5	.440
6	.117
7	.747
8	.031

Verbal Responses

If SLSs based on explicitly metacognitive understanding were being employed, we would expect participants to choose the same basket as the informed demonstrator in addition to providing a verbal explanation as to why they made this choice. However, we acknowledge that an individual (most likely a child) may use an understanding-based SLS without having the verbal ability to explain their understanding in which case there would be no verbal response to substantiate their basket choice. Nevertheless, we analysed verbal responses to the question: "Why did you choose that basket?" which was asked immediately after children/adults chose one of the two baskets. Verbal responses were placed into six categories as follows (Table 4.5A – Condition 1, Table 4.5B – Condition 2): one – explicitly metacognitive (copying the informed demonstrator due to their additional knowledge), two – majority bias (copying the uninformed demonstrators because they all selected the same basket), three – shifting/statistical (selecting the correct basket because it was chosen by the least number of demonstrators/least number of times), four – other reasoned (a seemingly reasoned response but one in which the reasoning was incorrect, inappropriate or based on an incorrect/misremembered fact), five – unreasoned (a response was given but it was incorrect/nonsensical) and six – no response (no verbal response was provided). For the first three of the above categories we also coded whether or not the verbal response corresponded to the basket choice made i.e. for Categories 1 (explicitly metacognitive) and 3 (shifting/statistical) did participants choose the same basket as the informed demonstrator

(the correct basket) and for category two (majority bias) did they choose the same basket as the uninformed demonstrators (the incorrect basket). We deemed this alignment to be an indicator that an individual was providing a true verbal justification for their basket choice, rather than producing a post-hoc justification for a choice they had later determined should have been made. We used this data to calculate the total percentage of participants (children of each age group and adults) who made a verbal response which corresponded to their basket choice (Tables 4.5A and 4.5B – Total matched basket choice).

Table 4.5

Summary of The Six Verbal Response Types (1-6) to the question: "Why did you choose that basket?". This includes the Number/Percentage of Participants in Conditions 1 (Table 4.5A, Adult and Child Data) and 2 (Table 4.5B, Child Data Only) Who Made Each Type of Response at Each Age (3-8 and Adult). For Response Types 1-3 We Also Display Whether Participants Chose a Basket Which Aligned With Their Verbal Response and the "Total Matched Basket Choice" Column Displays the Sum of These Columns at Each Age

A.

Age (whole years, number of participants in bracket)	Condition 1									Total matched basket choice
	Verbal response types 1-6 – number of participants									
	(% of participants in brackets)									
	1. Metacognitive total	Metacognitive & aligned (correct) basket choice	2. Majority bias total	Majority bias & aligned (incorrect) basket choice	3. Shifting/ statistical total	Shifting/ statistical & aligned (correct) basket choice	4. Other, reasoned	5. Unreasoned response	6. No response	
3 (13)	1 (7.7)	1 (7.7)	0	0	0	0	0	10 (76.9)	2 (15.4)	1 (7.7)
4 (16)	0	0	0	0	0	0	1 (6.3)	14 (87.5)	1 (6.3)	0
5 (15)	2 (13.3)	2 (13.3)	1 (6.7)	1 (6.7)	0	0	1 (6.7)	9 (60.0)	2 (13.3)	3 (20.0)
6 (13)	0	0	3 (23.1)	3 (23.1)	0	0	2 (15.4)	8 (61.5)	0	3 (23.1)
7 (16)	1 (6.3)	1 (6.3)	4 (25.0)	2 (12.5)	0	0	0	11 (68.8)	0	3 (18.8)
8 (14)	0	0	1 (7.1)	1 (7.1)	3 (21.4)	3 (21.4)	2 (14.3)	7 (50.0)	1 (7.1)	4 (28.6)
Adult (20)	12 (60.0)	10 (50.0)	3 (15.0)	3 (15.0)	2 (10.0)	2 (10.0)	2 (10.0)	1 (5.0)	0	15 (75.0)

B.

Age (whole years, number of participants in bracket)	Condition 2									Total matched basket choice
	Verbal response types 1-6 – number of participants (% of participants in brackets)									
	1. Metacognitive total	Metacognitive & aligned (correct) basket choice	2. Majority bias total	Majority bias & aligned (incorrect) basket choice	3. Shifting/ statistical total	Shifting/ statistical & aligned (correct) basket choice	4. Other, reasoned	5. Unreasoned response	6. No response	
3 (13)	0	0	0	0	0	0	0	11 (84.6)	2 (15.4)	0
4 (17)	2 (11.8)	1 (5.9)	1 (5.9)	1 (5.9)	0	0	1 (5.9)	13 (76.5)	0	2 (11.8)
5 (14)	1 (7.1)	1 (7.1)	0	0	0	0	3 (21.4)	10 (71.4)	0	1 (7.1)
6 (14)	1 (7.1)	0	0	0	0	0	2 (14.3)	11 (78.6)	0	0
7 (14)	1 (7.1)	0	1 (7.1)	1 (7.1)	1 (7.1)	1 (7.1)	1 (7.1)	9 (64.3)	1 (7.1)	2 (14.3)
8 (14)	1 (7.1)	1 (7.1)	3 (21.4)	3 (21.4)	2 (14.3)	1 (7.1)	3 (21.4)	5 (35.7)	0	5 (35.7)

We were especially interested in the proportion of participants who made the correct basket choice in addition to giving an “explicitly metacognitive” verbal response i.e. a response showing that they had used their understanding of the informed demonstrator’s mind to help them achieve task success. Although some individuals who chose a basket according to majority bias were able to explain the fact that they made this choice, we would not class these individuals as using an understanding-based SLS which utilised explicit metacognition because they were essentially stating that they acted according to the heuristic to “copy-the-majority” which did not enable task success. We designed our experiment so that use of knowledge attribution *only* should ensure that the correct basket was chosen, however we did not consider that what we will call the “shifting/statistical” strategy would also allow task success. This strategy was used by a very small number of participants who seemed to choose the correct basket based on the statistical reasoning that if they were incorrect then some chicks would still be found in the eggs removed by the three uninformed demonstrators (and the informed demonstrator’s first selection in Condition 1) and therefore that selecting the same basket as the informed demonstrator (that is, the last two locations selected in Condition 1 and the location selected three times in Condition 2) was the best way to ensure that the largest number of chicks were saved. Although an ability to verbalise use of this strategy showed that a social learning strategy based on some situational understanding was being employed, using this strategy required available information on the informed demonstrator’s additional knowledge to be overlooked and therefore we did not class this as an understanding-based, metacognitive SLS according to the definition presented in our introduction. The shifting strategy was placed into the same category as the statistical strategy because, as in the statistical strategy, the verbal response given focussed upon choosing the basket selected by fewer individuals (shifting away from the majority response), although it did not include an explanation as to why this might be beneficial.

Overall, children’s verbal responses did not indicate the use of explicitly metacognitive, understanding-based SLSs at any of the age groups tested in either Condition 1 or 2. Only a small proportion of children, between zero and 13% in each age group (0-2 children), chose the correct basket *and* provided an explicitly metacognitive verbal response (Tables 4.5A and 4.5B – Metacognitive & aligned (correct) basket choice). Only three additional children (ages 4, 6 and 7, Condition 2) provided an explicitly metacognitive verbal response in the absence of a correct basket choice (Tables 4.5A and 4.5B – Metacognitive total). Furthermore, in the subset consisting only of children who passed all the memory questions, not a single 8-year-old (the only age group found to

choose the correct basket at above chance levels in Condition 1) gave a verbal response indicative of explicit metacognition. The predominant verbal response type for the full dataset was “unreasoned response” which was given by at least 50% of children in each age group.

Contrastingly, 50% of adults (Condition 1 only) provided an explicitly metacognitive verbal response and a corresponding correct basket choice (Table 4.5A – Metacognitive & aligned (correct) basket choice). An additional 10% of adults gave such a verbal response but did not respond by choosing the correct basket, indicating that the verbal response may have been an afterthought for these adults. Therefore, a total of 60% of adults gave an explicitly metacognitive response (Table 4.5B – Metacognitive total), meaning that, as expected, this was the predominant verbal response type.

Discussion

We did not find evidence for the proposed shift from nonhuman-like (associative) to adult-like (understanding-based, explicitly metacognitive) social learning strategies during development. Identifying such a shift would support the hypothesis that humans use social information differently and that this may be related to developments in explicit metacognition (Dunstone & Caldwell, 2018), which are known to occur within the age range tested. Across our full sample of children (ages 3-8) the propensity to choose the correct/incorrect basket did not differ with age and the same, correct hiding location as the informed demonstrator was not chosen at above chance levels (Table 4.2 & Figure 4.2). Furthermore, those children who did make the correct choice were generally not able to support this action with a verbal justification indicative of explicitly metacognitive understanding (Tables 4.5A and 4.5B). When we excluded children who did not pass our three memory questions from the main sample (Table 4.3), the propensity to choose the correct/incorrect basket differed with age – only at age 8 were children more likely to choose the correct than the incorrect basket (and 8-year-olds were the only age group to choose the correct basket at above chance levels). However, this result was no longer significant when a Bonferroni correction was applied. Therefore, when considered alongside the lack of an explicitly metacognitive verbal response, it seems unlikely that this result at age 8 represents the proposed shift in use of social information.

We had expected that the majority of our adult participants (Condition 1 only) would select the correct basket. However, as with the child sample, adults did not choose the correct basket significantly above chance levels either as a complete sample or when we excluded participants who did not get all three memory questions correct. Unfortunately, this limited what we could interpret from the child results because we had aimed to compare this to successful performance in adults. We were interested in when children might begin to use

adult-like SLSs therefore if the task did not allow adults to demonstrate such performance it was unlikely to allow children to do so.

Most adults (83%) who chose the correct basket also provided a verbal explanation consistent with the use of an understanding-based, explicitly metacognitive social learning strategy (see Table 4.1 [Number participants who chose correct/incorrect basket – Condition 1, Correct] and Table 4.5A [Metacognitive & aligned (correct) basket choice]) whereas very few children who chose the correct basket provided such a response (<29% in any age category). This indicates that even though adults did not choose the correct basket above chance levels, when they did choose this basket, they were likely to be using an explicitly metacognitive, understanding-based SLS. Of course, it is possible that most children who chose the correct basket also used such a SLS. However, without a co-ordinating verbal response we cannot confirm this. It would therefore be extremely informative if future studies were designed to allow differentiation as to whether each participant's response was due to use of an explicitly metacognitive, understanding-based SLS or not, without needing to rely on verbal responses.

Adults were assumed to have explicitly metacognitive understanding of the demonstrators' minds and therefore to have no problem identifying that copying the informed demonstrator would lead to task success. Their poor performance thus indicates that many adults did not interpret the task as intended. This is supported by the finding that, although adults did not select the correct basket at above chance levels, all but one adult (5%) gave a reasoned response as to why they selected their chosen basket. Moreover, most adults who did interpret the task as intended (i.e. gave an explicitly metacognitive verbal response) selected the correct basket (see Table 4.5A – Metacognitive total and Metacognitive & aligned (correct) basket choice). It is possible that our experiment allowed for too much interpretation as to the knowledge state of the four demonstrators. For example, adults who interpreted the task incorrectly may have thought the demonstrators had knowledge outwith that they sought during the experiment e.g. that the uninformed demonstrators were all choosing the same basket because they already knew that it contained the chicks. Indeed, in further discussion some adults made comments to this effect or similar. Incorrect interpretations such as this, in conjunction with possible difficulties in remembering the exact actions of the informed/uninformed demonstrators, may have led to reasoning based on misinterpretations/inaccuracies. We therefore also calculated the total percentage of participants who made a basket choice (correct or incorrect) and gave a verbal response which matched (Table 4.5A – Total matched basket choice): 75% of adults and only of 29% children (mean for ages 3-8). This indicates that the majority of adults appeared to know, or at least were able to state, why they made the choice they did, even though as a group the

correct basket was not chosen at above chance levels.

We designed our task to ensure that the information provided by the informed and uninformed demonstrators was as accessible as possible. Yet for a substantial proportion of participants, remembering the actions of the informed/uninformed demonstrators appeared to be taxing: only 56% of children (and 65% of adults) in Condition 1 got all our memory questions correct, and 65% in Condition 2 (Table 4.3). As previously stated, when we included only those who got all the memory questions correct in our analyses, both children and adults were still not more likely to select the correct basket than the incorrect and did not do so above chance levels (other than the spurious result for 8-year-olds). This suggests that, although difficulty remembering the demonstrators' actions may have diminished participants' ability to use the available information, this factor alone was not likely to be responsible for the failure to detect a shift in SLS.

We acknowledge that our study was rather exploratory in nature. As far as we are aware, we are the first to directly test the above hypothesis, and (as discussed above in relation to adult performance) our study may not have been sensitive enough to allow detection of the proposed shift. In order to utilise explicit metacognitive ability participants needed to notice and remember that the informed demonstrator was privy to additional, task-relevant information and that information from this demonstrator should be used in preference to that from the uninformed demonstrators. It is possible that the cognitive load required to follow and weigh-up the conflicting information presented by the informed and uninformed demonstrators may have prevented use of the informed demonstrator's knowledge, even in children who had developed the explicitly metacognitive ability required to understand that this demonstrator was more informed. Therefore, had our study included older children (e.g. aged 9+), these children may have been better able to utilise information from the informed demonstrator due to further developments in executive functions throughout middle childhood and into adolescence (Best & Miller, 2010). However, in light of the poor adult performance, this is impossible to determine without making additional changes to the task (discussed below).

We have discussed how use of explicit metacognition to devise an understanding-based SLS in this task would have required an ability to over-ride copying the majority of demonstrators. In addition, children may have had to overcome an implicit "feeling of knowing" which basket contained the chicks. False feelings of knowledge such as this have been proposed to occur in young children who have been exposed to the potential contents of containers in partial exposure knowledge-assessment tasks (Kloo et al., 2017; Rohwer et al., 2012; Sodian & Wimmer, 1987). Young children likely over-estimate their knowledge in these tasks/fail to state their ignorance because they have difficulty understanding that

knowledge comes from a particular source until around age 6 (Kloo et al., 2017; Rohwer et al., 2012; Sodian & Wimmer, 1987), and possibly older depending on the task. Our task could be considered to fall into the category of “partial exposure” (although this is debatable) because at the start of the task children were shown chicks and yolks and told that the chicks were hidden inside the eggs in one container, and the yolks inside eggs in the other container. However, as for the points discussed above, the poor adult performance makes determining the impact of partial exposure on children’s performance difficult i.e. any age effects may have been masked by incorrect interpretation of the task leading to children with explicitly metacognitive understanding employing a range of strategies.

In order to address our main research question, future tasks would need to allow adults to perform as expected and therefore detection of the proposed shift in social information use. As discussed above, it seems that adults did not interpret the task as intended and so any future task would need to ensure that participants with explicitly metacognitive understanding could more readily establish the rule which would lead to task success. Future designs may therefore benefit from removal of a need to remember the actions of as many demonstrators and to over-ride use of a default strategy such as copy-the-majority. Instead, it may be informative (at least in the first instance) to concentrate on detecting whether one’s ability to copy one informed, rather than one uninformed, demonstrator increases in line with developments in explicit metacognition. In such a context, we would expect it to be more obvious that such a strategy allows for task success. Of course, this approach may not be strictly ecologically valid (over-riding a default bias may be needed for an individual to devise an understanding-based SLS in most real-world situations), but if we are to determine if explicitly metacognitive capacities are important for differential use of social information then this may be a necessary first step.

Heyes (2016, 2018) defines explicitly metacognitive SLSs as “consciously represented, reportable, domain-specific rules” which differ from other explicit metacognitive rules because they “specify conditions in which it is advisable to engage in social rather than asocial learning” and “from whom one should learn”. By this definition both our “metacognitive” and “majority bias” (Table 4.5A) verbal responses would be classed as explicitly metacognitive SLSs. 65% of adults and below 29% of children in each age category would be classed as responding according to one of these strategies whilst choosing the corresponding basket (i.e. correct for a “metacognitive” response and incorrect for a “majority bias” response – see Table 4.5A). However, according to our definition of explicitly metacognitive, understanding-based SLSs this would drop to 50% of adults and below 13% (or less, depending on age group) of children (Table 4.5A – Metacognitive & aligned (correct) basket choice). Participants explicitly stating that they were copying the

majority of uninformed demonstrators would not be included because they would not be displaying evidence of devising a strategy based on explicitly metacognitive understanding of the specific task situation (i.e. the informed demonstrator's additional knowledge). However, whichever definition is used, the absence of a verbal response does not necessarily mean that explicitly metacognitive rules were not being represented in the minds of children/adults. Verbal responses can therefore not be used to definitively differentiate between child and adult use of such rules in tasks like ours.

If future studies confirm that adults utilise social learning strategies distinct from those of nonhumans then a further interesting research question could address whether some of these strategies are: explicitly metacognitive, understanding-based SLSs as per our definition; explicitly metacognitive SLSs as defined by Heyes (2016); both; or neither. This would require differentiating a strategy which has been devised by the individual from one which the individual has learnt to use through social interaction. Presumably the individual using either strategy would have an awareness of it and an ability to report it. However, the strategies could be differentiated through determining whether an individual was able to verbalise the strategy/rule used (i.e. their reason for making a particular choice) *in addition to* the reasoning underlying their use of the strategy/rule (as per our definition), or just the strategy/rule used (satisfying the definition of an explicitly metacognitive SLS provided by Heyes (2016)).

To conclude, we found no evidence to support the theory that humans use social information more selectively due to developments in explicit metacognition (Dunstone & Caldwell, 2018). We had aimed to identify whether a shift from nonhuman-like to adult-like social learning strategies occurs as children develop explicit metacognition. However, the poor performance of adults, whom we assumed to have developed an ability for explicit metacognition, limits the conclusions which can be drawn from the present study. It appears likely that adults did not interpret the task as we intended. Therefore, in order to address our research questions, the design of any future studies would need to ensure that adults/children with explicitly metacognitive understanding perform as expected according to the task paradigm.

Chapter 5: Who Knows? Does Children's Developing Understanding of Others' Minds Affect Their Social Learning Strategies? Part 2.

Introduction

In the previous study (Chapter 4) we examined whether there is a shift in the way in which social information is used as understanding of the mind goes from the level of a human infant to that of an adult. We hypothesised that at some point in development children may gain the ability to employ social learning strategies based on their explicit metacognitive understanding (in this case, of another agent's knowledge) rather than using heuristic-based social learning strategies (SLSs) which do not require such understanding. The adult-like SLSs were referred to as understanding-based, explicitly metacognitive SLSs and we did not find evidence for a shift in strategy use consistent with the use of these SLSs in the age range tested (3-8 years). However, the performance of an adult control group was also poor, and adults did not use the social information as we expected based on their explicitly metacognitive understanding. It therefore appears that our task was not measuring use of understanding-based, explicitly metacognitive SLSs as intended, which limited what could be interpreted from the child results.

In our previous task an informed demonstrator had privileged, visual access to the location of a hidden reward whereas three uninformed demonstrators did not. In order for this task to detect the aforementioned shift in social information use, it was important that participants with an explicitly metacognitive understanding of the informed demonstrator's mind were able to devise a strategy to copy this demonstrator and achieve task success. However, adults (presumably possessing the required explicit understanding) did not copy the informed demonstrator at above chance levels. Adult verbal responses indicated that many adults did not interpret the task as intended and adults employed a range of strategies, leading to variable performance. It thus appears that the "correct" strategy was not obvious enough, even to those who did have the ability to employ explicitly metacognitive SLSs. In this follow-up study we therefore set out to increase the transparency of the correct strategy so that adults (and children who had developed the required understanding) would interpret the task as we had originally expected. It was important that adults achieved this benchmark, to which child performance could be compared, and this was a fundamental goal of the current study. As in the previous study, we also aimed to determine if there is a shift in social information use during ages 3-8 (discussed further below).

In order to achieve success in our previous task, in addition to using explicitly metacognitive understanding, children needed to continuously monitor and remember the choices of four different puppet demonstrators (three uninformed, one informed). The results

of our memory questions revealed that a large proportion of participants (including many children aged 3 to 8 and some adults) found this difficult and, although we do not believe memory was the only factor limiting task success, we aimed to reduce this difficulty in the current follow-up study. In the previous study the use of a minimum of three demonstrators (although we used four) was necessary because we were assessing an expected shift from use of an heuristic strategy to copy a group of uninformed demonstrators to use of a strategy based on the explicitly metacognitive understanding that the one informed demonstrator had additional, task-relevant knowledge. It is therefore possible that the executive function requirements of the task prevented children from using their explicitly metacognitive understanding of the informed demonstrator's knowledge to succeed. However, as noted above, we were unable to draw firm conclusions from the child results because adults also struggled with the task.

In this current study we thus aimed to lessen the need to remember and calibrate multiple demonstrations in order to reduce both the cognitive load and the potential for task misinterpretation (as discussed above). In regard to task misinterpretation – the additional demonstrators and need to pit one strategy against another (see Chapter 4), may have masked the correct strategy and led a number of adults to make the incorrect choice. The current task instead focused on determining when in development children will use information from one informed individual in preference to just one uninformed individual. This differed from the research question presented in the former study because we were no longer looking to identify a potential shift from use of a default strategy to use of an understanding-based, self-derived social learning strategy. Instead, we were purely interested in identifying the possible developmental stage at which deriving an understanding-based, explicitly metacognitive SLS, based on understanding of another individual's mind, becomes possible (or at least becomes possible given our task demands).

As in our previous experiment, in the current study we aimed to identify any potential shift in social information use with age (3-8). If the age at which such a shift occurs aligns with the age at which developments in explicit metacognition (e.g. understanding of false belief, how evidence relates to knowledge and/or partial knowledge exposure – background discussed in the introduction to the previous study in Chapter 4) have been documented, then this would indicate that these developments may be involved in determining approach to social information use. That is, developments in explicit metacognition, considered unique to humans, may be required in order to derive an understanding-based, explicitly metacognitive SLS.

The informed individual in this current study again had visual perceptual access to task-relevant knowledge (the contents of hiding locations), therefore it was possible to use

understanding of this individual's mind, specifically that they had knowledge due to visual access, to determine that they alone should be copied in order to locate the hidden objects. We hypothesised that the aforementioned changes, to decrease cognitive load and limit the potential for task misinterpretation, would make it easier for both children and adults to use the information from this demonstrator. In fact, we predicted that we might observe an increase in the number of children choosing the same, correct hiding location as the informed demonstrator in line with known developments in explicit metacognition (e.g. understanding of false belief, how evidence relates to knowledge and/or partial knowledge exposure). We again had no a priori expectations regarding an exact developmental stage at which children might begin to prioritise information from the informed demonstrator because we do not know which explicit metacognitive developments might be most critical.

In this current task we tested children aged 3-8 years and adults. The specific task goal presented to all participants, and the informed/uninformed individuals (puppets), was to find objects hidden inside plastic eggs in one of four locations (baskets). The uninformed individual (without visual access to the objects hidden inside the eggs) was observed choosing an incorrect location, but the informed individual was observed gaining additional task-relevant perceptual information (opening an egg from each basket) before choosing the correct location and could therefore be considered to know in which location the objects were hidden. After watching the informed/uninformed individuals each select a basket, participants were required to choose one of the four baskets themselves, that which they believed contained the hidden objects. They were not made aware of whether the informed/uninformed individuals had correctly chosen the location containing the desired objects until the end of the experiment because we did not want children to base any strategy on visible success/failure of the models, but rather on their interpretation of the models' actions.

In this second experiment the number of location choices (baskets) was increased from two (see Chapter 4, Figure 4.1) to four (Figure 5.1A) so that there were two baskets not chosen by either demonstrator. In the previous study children had to make a binary choice as to which basket contained the hidden objects and this made it difficult to distinguish between those who were choosing at random and those possibly following alternative strategies. Increasing the number of potential hiding locations thus enabled us to differentiate between children who did not copy the informed demonstrator, but did still copy one of the demonstrators, from those who chose a basket which wasn't selected by either demonstrator. This differentiation would have been possible with three baskets, but we included a fourth so that there was an equal number of baskets either chosen by one of the two demonstrators, or not.

Figure 5.1

Experimental Materials Including the Following: Four Plastic Baskets Each Containing Twelve Eggs (A), an Egg With a Black Dot at the Apex (B), Example Eggs Containing a Yolk and a Chick (C), Cow and Dog Puppets (D), Pig and Duck Puppets (E), and the Hen Puppet (F)



Method

Participants

131 children aged 3 to 8 were recruited from the following UK locations: a private nursery and two Scout Groups (Beavers – for girls and boys aged 6-8) in Stirling, and Glasgow Science Centre (a science museum with an entry fee) in Glasgow. Twelve children were excluded for the following reasons: experimenter error (n=2, aged 3 and 6), parental assistance with the task (n=3, two aged 3, one aged 4), lack of interest in the task (n=1, aged 8), it was not possible to communicate the task instructions to the child (n=2; aged 3 and 8), incomplete child details on the consent form (n=1, aged 5), unresponsive (n=1, aged 4) and child interference with the task equipment (n=2, aged 6 and 7). The final sample consisted of 119 children aged 3 to 8 (M= 71 months, range= 36 – 107, SD= 21, 50 male and 69 female). There were between 19 and 21 children in each age group (Table 5.1). 18 adults aged 21 – 69 were recruited from the University of Stirling via advertisement on the University intranet and none were excluded (M= 422 months, range= 253 – 832, SD= 204, four male and 14 female).

Table 5.1

Total Number of Participants (Age 3-8 and Adult) and Number Who Chose the Informed, Uninformed or One of the Other Baskets in Trials 1 and 2. The Percentage of Total Participants is Shown in Brackets to One Decimal Place

Age (whole years)	Total number participants	Number participants who chose informed/uninformed/other baskets (% of total participants in brackets)					
		Trial 1			Trial 2		
		Informed	Uninformed	Other	Informed	Uninformed	Other
3	20	3 (15.0)	6 (30.0)	11 (55.0)	6 (30.0)	8 (40.0)	6 (30.0)
4	21	8 (38.1)	4 (19.0)	9 (42.9)	4 (19.0)	8 (38.1)	9 (42.9)
5	20	5 (25.0)	5 (25.0)	10 (50.0)	6 (30.0)	6 (30.0)	8 (40.0)
6	19	4 (21.1)	5 (26.3)	10 (52.6)	9 (47.4)	4 (21.1)	6 (31.6)
7	20	5 (25.0)	3 (15.0)	12 (60.0)	5 (25.0)	6 (30.0)	9 (45.0)
8	19	6 (31.6)	2 (10.5)	11 (57.9)	5 (26.3)	8 (42.1)	6 (31.6)
adult	18	16 (88.9)	2 (11.1)	0	18 (100.0)	0	0

Ethics statement: This research was approved by the University of Stirling, General University Ethics Panel (reference: GUEP 374A). Written, informed consent was obtained from all adults and the parent or guardian of all children prior to their participation. Children were asked if they would like to participate, were continuously monitored for assent and were rewarded with a sticker regardless of task completion. Adults were provided with monetary compensation for their time and were informed that they could stop participating at any point during the experiment.

Materials

Large farmyard animal puppets – a cow, dog, duck, hen and pig (The Puppet Company©) were used as demonstrators in social information trials (Figures 5.1D-F). The hen always acted in a fixed role, and we counterbalanced whether the remaining four puppets acted as informed/uninformed demonstrators. The choice stimuli were four yellow and four green, plastic baskets (base: 110mm x 110mm, top opening 150mm x 150mm, height: 140mm) for use in Trial 1 and Trial 2 respectively (Figure 5.1A). Each basket contained twelve gold, plastic eggs (diameter centre: 40mm, height: 60mm). One basket of each colour contained eggs with imitation chicks (height: 30mm, width: 15mm, length: 15mm) inside, and the other baskets contained eggs with imitation yolks inside. These were made from balls of yellow tissue paper of the same size and weight as the chicks (Figure 5.1C). Within each colour set (yellow or green), each basket was assigned round (diameter: 9 mm) stickers, either red, pink, brown or blue, to differentiate it from the remaining three baskets in the colour set (Figure 5.1A). Four stickers of the same colour were placed onto each basket, two on the front and two on the back. One of the twelve eggs in each basket was empty and marked with a small black dot at the apex (Figure 5.1B).

Procedure

At Glasgow Science Centre child testing was carried out in a public space, separated from other visitors by a table. At the nurseries and Beavers groups testing took place in a quiet room close to the main playroom. At Glasgow Science Centre only, less confident children were accompanied in the testing area by a parent or guardian, who was instructed not to provide the child with any task-relevant assistance. In both locations the experimental task was carried out on a table-top with the child seated opposite the experimenter (operating the puppets) and next to the research assistant (live coding the child's responses). A verbal script was used by both the experimenter and the research assistant (see Appendix "Experimenter Verbal Scripts"). Prior to starting the experiment, the four

yellow baskets, to be used in Trial 1, were positioned in a straight line across the centre of the table (Figure 5.1A). From the child's perspective the order of the baskets, from left to right and according to sticker colour, was blue, brown, pink and red. Eleven eggs were placed into each basket, but only one of the baskets (that to be selected by the "informed" demonstrator – see below) contained the eggs with the chicks inside. An empty egg with a black dot at its apex (Figure 5.1B) was then placed into the top of each basket. The set-up for Trial 2 (including the order of the sticker colours on the four baskets) was identical except that green baskets were used and the eggs with chicks inside resided in a different basket. The testing procedure for adults was identical to that for children except that adults were tested individually, in a small room at the University of Stirling; it was explained to them that the task was designed to be engaging for young children.

Prior to Trial 1, all participants were introduced to the task through a short interaction with the hen puppet. Child participants were asked if they would like to play a game (see Appendix "Experimenter Verbal Scripts" for exact language used throughout) and adults were asked whether they would like to take part in a task we were running with children. Participants were told that the hen had two types of eggs which looked the same on the outside – some with chicks inside and some with yolks inside "like the eggs we eat". They were shown examples of these eggs and invited to look at/touch the contents. For Trial 1, participants were then informed that the silly hen had put her chicks into one of the four yellow baskets, but no longer knew which one and had therefore lost her chicks. They were asked: "Can you help me find out which basket has my baby chicks inside?" Finally, participants were told that the hen would go and find some animals (cow, dog, duck and pig puppets) to help find her chicks. A social information trial followed by a test trial was then conducted for Trial 1 and, subsequently, for Trial 2.

Social Information Trials. In Trial 1, the four yellow baskets were searched by the puppets – Cow (first) and Dog (second), and in Trial 2 the four green baskets by Pig (first) and Duck (second). The exact actions performed during the search depended on whether the puppet was deemed to be an informed or uninformed demonstrator; one puppet in each trial acted as each type of demonstrator. The puppets always searched in the same, fixed order, but we counterbalanced whether the cow or dog acted as the informed demonstrator in Trial 1. This then determined whether the informed demonstrator was the first or second demonstrator in Trial 2 because we reversed the order e.g. informed, uninformed in Trial 1 would be matched with uninformed, informed in Trial 2. We assigned half the participant numbers within each age group (3 to 8 and adults) to the order: Trial 1 informed, uninformed and half to the order: Trial 1 uninformed, informed.

A demonstrator categorised as uninformed would choose one of the four baskets without having visual access to the contents of the eggs in any basket, whereas an informed demonstrator would peek inside an egg from each basket, and therefore be privy to its contents, before making a choice. The uninformed demonstrator therefore hovered over the baskets in turn, stopped to look inside, and at the last (pre-determined) basket stated: "I choose this one!" "Hen should keep this basket." This was accompanied by the demonstrator grabbing and shaking the tub slightly to aid the child's memory of the selection. The informed demonstrator performed the same actions, and made the same verbal comment, with the addition of peeking inside an egg from each of the baskets. This involved opening the egg slightly and holding the opening close to the eye. To ensure that these actions did not allow the participant to see the contents of the egg, the informed demonstrator opened the one empty egg from each tub (identifiable by a black dot) very carefully so that it was evident that the egg was being opened but not that it was empty. To direct a participant's attention towards the demonstrators as they were looking inside the tubs/eggs, the research assistant stated: "Look at what [Cow/Dog/Duck/Pig] is doing!"

For each participant, we assigned two pairs of colours randomly generated from the four sticker colours. The two colours in Pair 1 and 2 determined the first and second demonstrators' final colour choices in Trial 1 and 2 respectively i.e. the last basket each demonstrator interacted with and thus their chosen basket. The order of the three baskets each demonstrator selected prior to this was kept as varied and balanced as possible by the experimenter making the selections (ensuring that the demonstrators did not search the baskets in order from left to right, or similar), however the exact order was not predetermined. The chicks were always placed inside the basket which was to be chosen by the informed demonstrator and the demonstrator always selected this location last.

Test Trials. Following a social information trial, the research assistant gave the participant the following instructions: "Now you can choose which basket you think we should give to Hen. You should choose the basket that you think has Hen's baby chicks inside because Hen is looking for her baby chicks." The assistant then asked: "Which basket should we give to Hen?" Both verbal and non-verbal responses (i.e. pointing at one of the baskets) were accepted, and the participant was told that they would be able to look inside their chosen basket shortly. The assistant then asked the following questions: "Why did you choose this basket?" "What do you think is inside this basket?" This last question was only asked if the participant had not included the answer as part of their answer to the previous question (and therefore not necessary for most adults). Participants were asked a further two questions to assess their memory of the social information trials. These were as follows (for

the puppets presented in Trial 1 as an example): “Which basket did Cow choose?” “Which basket did Dog choose?” Whilst these memory questions were being asked, the experimenter held up the relevant puppet to aid memory. These questions were always asked in the same order (the order in which the puppets were presented during the trial), but which puppet acted as the informed demonstrator was counterbalanced. The responses to all questions (verbal or non-verbal, where appropriate) were live coded by the research assistant.

Once the social information and test trials for Trial 1 and then 2 were complete, participants were invited to look inside the baskets they had selected in Trial 1 and 2 in order to discover whether they had found the chicks. If the participant failed to find the chicks then they were asked if they would like to look inside the remaining baskets. Total testing duration was approximately 10 minutes for children and 5-10 minutes for adults, and the interval between Trials 1 and 2 was kept as short as possible. Following Trial 1, but before Trial 2, the yellow basket chosen by the participant was placed at one end of the table (but within sight) and the remaining baskets were removed and replaced with the green baskets.

Results

All generalised linear mixed effects models (GLMM) and generalised linear models (GLM) were carried out with either the logit link (binomial data, family = binomial) or log link (count data, family = poisson) and the lme4 package (lmer and glm functions for GLMMs and GLMs, respectively) using R (R Core Team, 2018). Our default choice for the random effects structure for each model included by-participant random slopes for variables which varied within participant, following Barr et al. (2013), to keep random effects structures “maximal” where possible. Where the maximal model resulted in non-convergent or singular fit models, random slopes were removed followed by random intercepts where necessary until a convergent, non-singular model was obtained. p -values < .05 were taken as statistically significant, unless stated otherwise.

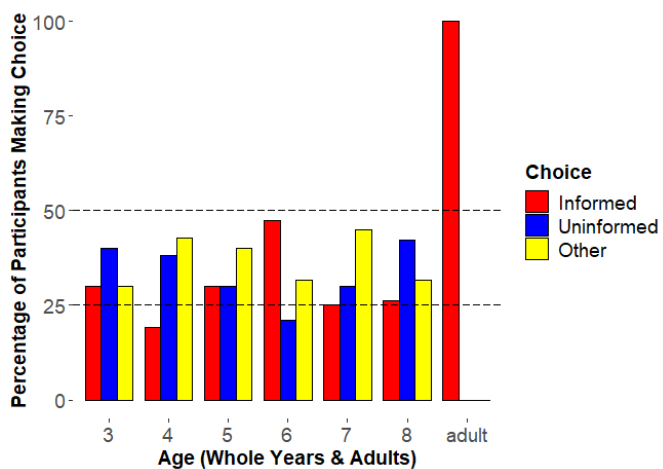
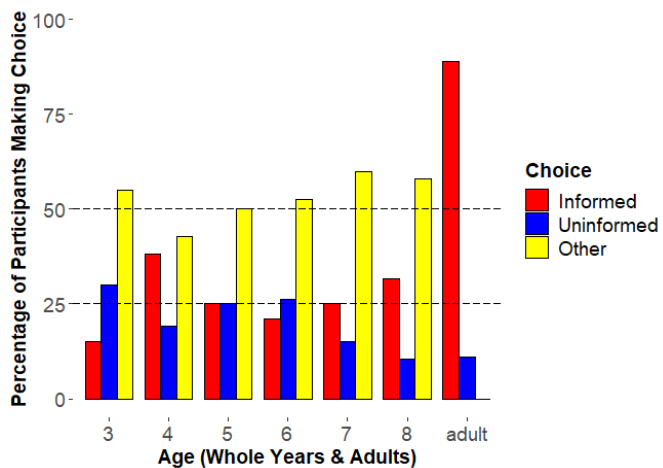
Basket Choices

We were primarily interested in whether there was a shift towards more adult-like, understanding-based task responses (i.e. copying the informed demonstrator’s correct choice) with increasing age (3-8 years) in Trials 1 and 2. We first examined whether there was any pattern in the adult and children’s responses consistent with biased responding or whether participants appeared to be choosing baskets at random. We were interested in comparing child performance to that of the adults, which we expected to be biased towards copying the informed demonstrator.

We calculated the number and percentage of participants who made the same (correct) choice as the informed demonstrator, the same (incorrect) choice as the uninformed demonstrator, or chose one of the two other (incorrect) baskets, for adults, children of each age group and trial (Table 5.1 & Figure 5.2). If participants chose one of the four baskets at random, we would expect 25 percent to have chosen the informed basket, 25 percent the uninformed, and 25 percent each of the other baskets. As expected, in adults, exact goodness of fit tests (Table 5.2), for Trials 1 and 2 both separately and combined, showed that the proportion who made the choice of informed, uninformed or other was significantly different to chance (p values $< .001$). Post-hoc, exact goodness of fit tests with Bonferroni correction and an alpha level of .017, again for Trials 1 and 2 both separately and combined, revealed that significantly more adults chose the informed basket ($p < .001$) and significantly fewer chose one of the other baskets ($p < .001$) than expected according to chance proportions. Significantly fewer adults choose the uninformed basket than expected by chance in Trial 2 ($p = .007$), and for Trials 1 and 2 combined ($p = .004$), but not in Trial 1 ($p = .192$). However, in children these three proportions were not significantly different to those expected by chance at any age group, in either Trials 1 or 2 separately or combined, (all p values ages 3-8 $> .05$). This supports the hypothesis that children may have been choosing baskets as if at random.

Figure 5.2

Percentage of Participants Who Chose the Informed, Uninformed or One of the Other Baskets by Age (Whole Years & Adults) in Trial 1 (Top Panel) and Trial 2 (Bottom Panel)



Note. The dashed lines show the percentage performance we would expect if participants were selecting the baskets at random – 25% informed, 25% uninformed and 50% other.

Table 5.2

Summary of Two-Sided, Exact Goodness of Fit Tests to Test Whether the Proportion of Participants Who Chose the Correct, Uninformed or One of the Other Baskets Was Different to the Expected Chance Proportions for Trials 1 and 2. Full Dataset

Trial number	Age (whole years)	<i>p</i> value – exact goodness of fit test (full data set)
1	3	.575
1	4	.410
1	5	.957
1	6	.957
1	7	.523
1	8	.373
1	adult	< .001
2	3	.187
2	4	.410
2	5	.733
2	6	.130
2	7	.957
2	8	.177
2	adult	< .001
Both	3	.381
Both	4	.658
Both	5	.856
Both	6	.472
Both	7	.938
Both	8	.831
Both	adult	< .001

In order to identify whether children shifted towards use of the correct strategy with age we scored participants as having chosen the informed/correct (1) or one of the three incorrect (0) baskets (one uninformed, two other) and analysed the effect of child age on “choice correct” (0 or 1) by performing a GLM (family = binomial) with choice correct as the dependent variable and age (in days, centred and scaled to measure thousands of days) as the fixed effect, for Trials 1 and 2 separately. There was no main effect of age on choice correct in either Trial 1 ($b = 0.132$, $SE = 0.335$, $Z = 0.395$, $p = .693$) or 2 ($b = 0.107$, $SE = 0.323$, $Z = 0.330$, $p = .741$). Furthermore, Fisher’s exact tests (two-sided) confirmed that the percentage of participants choosing the correct/uninformed/other baskets (Table 5.1) was independent of age category in both Trials 1 ($p = .845$) and 2 ($p = .840$). Children’s ability to

choose the informed basket and achieve task success therefore did not improve with age as would be expected if they were beginning to use more understanding-based strategies.

In our previous experiment, memory questions revealed that participants (in particular children but to some extent also adults) found it difficult to remember from which of the two baskets each of the four demonstrators removed an egg and whether each demonstrator was informed (peeked inside each of the three eggs retrieved), or uninformed (did not peek inside the retrieved eggs). Consequently, we simplified the design of the current study, as has been discussed previously, and again asked participants some memory questions. These were as follows (e.g. for the puppets presented in Trial 1): “Which basket did Cow choose?” “Which basket did Dog choose?” We were thus able to create subsets of our Trial 1 and 2 data which included only participants who passed both memory questions for the relevant trial. For the adults these subsets were almost identical to the complete adult data set, with only one or two adults, in Trials 2 and 1 respectively, failing to pass both of our memory questions (see Tables 5.1 and 5.3). In children the complete data set and these subsets were more variable (Tables 5.1 and 5.3). Therefore, a significantly higher proportion of correct basket responses for the subsets, and/or an age effect not present in the full child dataset, would indicate that difficulty with continuously monitoring/remembering the demonstrators’ selections may have impeded ability to make the correct choice as in our previous experiment. We therefore calculated the percentage of participants who chose the informed, uninformed or one of the other baskets for each age and trial for these subsets (Table 5.3).

Table 5.3

Total Number of Participants (Age 3-8 and Adult), in the Subsetted Data Containing Only Participants Who Passed Both Memory Questions, and Number Who Chose the Informed, Uninformed or One of the Other Baskets in Trials 1 and 2. The Percentage of Total Participants is Shown in Brackets to One Decimal Place

Age (whole years)	Total number participants (subsetted data)		Number participants who chose informed/uninformed/other baskets for the subsetted data containing only participants who passed both memory questions (% of total subsetted participants in brackets)					
			Trial 1			Trial 2		
	Trial		Informed	Uninformed	Other	Informed	Uninformed	Other
	1	2						
3	7	5	1 (14.3)	2 (28.6)	4 (57.1)	1 (20.0)	3 (60.0)	1 (20.0)
4	12	7	4 (33.3)	2 (16.7)	6 (50.0)	1 (14.3)	4 (57.1)	2 (28.6)
5	12	11	3 (25.0)	2 (16.7)	7 (58.3)	4 (36.4)	3 (27.3)	4 (36.4)
6	14	11	3 (21.4)	5 (35.7)	6 (42.9)	5 (45.5)	2 (18.2)	4 (36.4)
7	17	10	4 (23.5)	3 (17.6)	10 (58.8)	1 (10.0)	2 (20.0)	7 (70.0)
8	14	13	5 (35.7)	1 (07.1)	8 (57.1)	3 (23.1)	6 (46.2)	4 (30.8)
adult	16	17	14 (87.5)	2 (12.5)	0	17 (100.0)	0	0

As expected, adult performance was comparable to that from the (almost identical) full adult dataset – the relative proportions of the basket choices were again significantly different to chance in Trials 1 and 2 (p values $< .001$, Table 5.4). Moreover, post-hoc, exact goodness of fit tests with Bonferroni correction and an alpha level of .017, again for Trials 1 and 2, revealed that significantly more adults chose the informed basket ($p < .001$) and significantly fewer chose one of the other baskets ($p < .001$) than expected according to chance proportions. Significantly fewer adults choose the uninformed basket than expected by chance in Trial 2 ($p = .008$) but not in Trial 1 ($p = .277$). As in adults, we did not find evidence of differential performance in the child subsets compared to the full child dataset. Exact goodness of fit tests for the subsets (Table 5.4) again showed that the proportion of children who chose the correct, uninformed or other baskets was not significantly different to that expected by chance at any age group in either Trial 1 or 2 (all p values ages 3-8 $> .05$). GLMs for Trial 1 and 2 “all memory questions correct” child data were also performed as for the full child datasets and revealed no main effect of age on choice correct in either Trial 1 ($b = 0.174$, $SE = 0.447$, $Z = 0.390$, $p = .697$) or 2 ($b = -0.090$, $SE = 0.539$, $Z = -0.166$, $p = .868$). Fisher’s exact tests (two-sided) confirmed that the percentage of participants choosing the correct/uninformed/other baskets was independent of age category in both Trial 1 ($p = .910$) and Trial 2 ($p = .455$).

Table 5.4

Summary of Two-Sided, Exact Goodness of Fit Tests to Test Whether the Proportion of Participants Who Chose the Correct, Uninformed or One of the Other Baskets Was Different to the Expected Chance Proportions for Trials 1 and 2. Including Only Subsetted Data i.e. Participants Who Got all Our Memory Questions Correct

Trial number	Age (whole years)	<i>p</i> value – exact goodness of fit test (subsetted data)
1	3	.898
1	4	.703
1	5	.824
1	6	.778
1	7	.740
1	8	.225
1	adult	< .001
2	3	.551
2	4	.371
2	5	.622
2	6	.383
2	7	.366
2	8	.245
2	adult	< .001

In addition to scoring participants as making the correct (1) or incorrect (0) choice on each trial we generated a total score (0-2) variable which was the summed score for Trials 1 and 2 and therefore represented each participant's performance across the entire task. We expected most adults to achieve a score of 2. However, as with our choice correct (0 or 1) variable, in children we were interested in whether total score (0-2) increased with age (Table 5.5, Figure 5.3). We therefore performed a GLMM (family = poisson) with total score (0, 1 or 2) as the dependent variable and age (in days, centred and scaled to measure thousands of days) as the fixed effect. There was no main effect of age on total score ($b = 0.086$, $SE = 0.197$, $Z = 0.435$, $p = .664$).

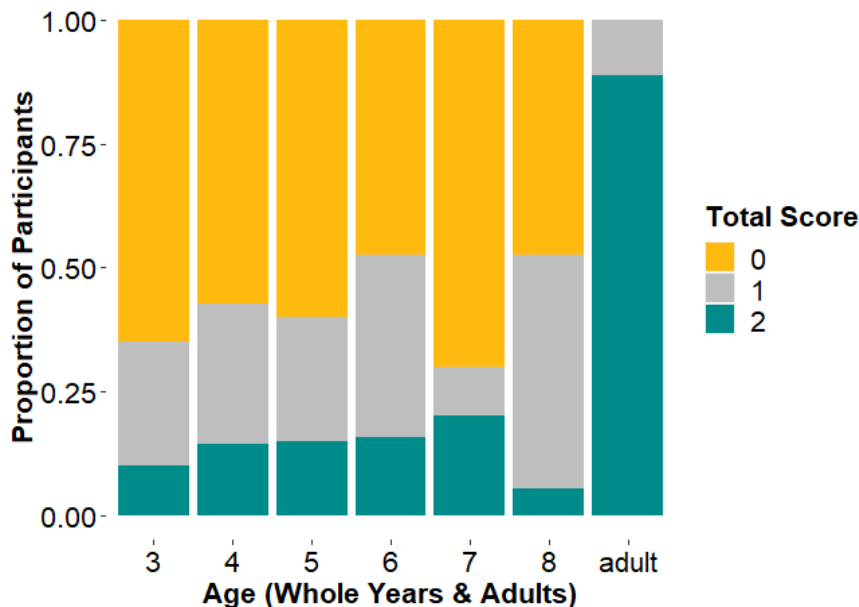
Table 5.5

Number of Participants (Aged 3-8 and Adult) With Total Scores of 0, 1 or 2 From Trials 1 and 2 Combined. The Percentage of Total Participants is Shown in Brackets to One Decimal Place

Age (whole years, number of participants in brackets)	Total score trials 1 & 2 – number of participants (% of total participants in brackets)		
	0	1	2
3 (20)	13 (65.0)	5 (25.0)	2 (10.0)
4 (21)	12 (57.1)	6 (28.6)	3 (14.3)
5 (20)	12 (60.0)	5 (25.0)	3 (15.0)
6 (19)	9 (47.4)	7 (36.8)	3 (15.8)
7 (20)	14 (70.0)	2 (10.0)	4 (20.0)
8 (19)	9 (47.4)	9 (47.4)	1 (05.3)
adult (18)	0 (00.0)	2 (11.0)	16 (88.9)

Figure 5.3

Proportion of Participants With a Total Score of 0, 1 or 2 From Trials 1 and 2 Combined by Age (Whole Years & Adults)



Most adults performed as expected, and 88.9% chose the correct basket in both trials. An exact goodness of fit test (Table 5.6) revealed that the proportion scoring 0, 1 or 2 was significantly different to that expected by chance ($p < .001$) and post-hoc, exact goodness of fit tests with Bonferroni correction and an alpha level of .017, showed that significantly more adults scored 2 ($p < .001$) and significantly fewer scored 0 ($p < .001$) than expected by chance. The number of adults scoring 1 was marginally significantly less than expected by chance ($p = .017$) if the alpha level under the Bonferroni correction was adhered to. Contrastingly, exact goodness of fit tests revealed that the proportion of children scoring 0, 1 or 2 was not significantly different to chance for any age group except age 7 (p values excluding age 7 $> .05$, p value age 7 = .005). At age 7 post-hoc, exact goodness of fit tests, with Bonferroni correction and an alpha level of .017, showed that significantly fewer children scored 1 ($p = .010$) than expected by chance. This may suggest a bimodal distribution between children of this age who perhaps understood what the correct strategy was, and thus scored 2, and those who did not, and thus scored 0. However, this result must be interpreted with caution because the proportion of children scoring 2 ($p = .033$) and 0 ($p = .263$) did not differ from chance (although if the Bonferroni correction was not used the proportion scoring 2 did of course differ from chance at the standard alpha level of .05).

Overall responding therefore did not appear to be consistent with the use of understanding-based SLSs at age 7, or indeed any child age group.

Table 5.6

Summary of Two-Sided, Exact Goodness of Fit Tests to Test Whether the Proportion of Participants Scoring 0, 1 or 2 Was Different to the Expected Chance Proportions

Age (whole years)	p value – exact goodness of fit test
3	.573
4	.465
5	.375
6	.494
7	.005
8	.720
adult	< .001

Thus far we have looked at children's basket choices across Trials 1 and 2 in isolation (informed, uninformed or other basket choice) and combination (total scores). However, we were also interested in how children's performance may have changed as they proceeded through the task from Trial 1 to Trial 2 because this could provide insight into how children approached the task and why they performed so poorly overall. Following the completion of Trial 1 children were not informed as to whether they had successfully found the hen's lost chicks and it was made clear that their chosen basket would be put aside until the end of the game. The baskets used in Trial 2 were a different colour to those in Trial 1, therefore children shouldn't have found it difficult to disassociate these and to understand that any information provided by the Trial 1 demonstrators (particularly in regards to selection of a basket with a particular coloured sticker) should be disregarded in Trial 2. We therefore expected children to approach the second trial as they did the first trial, although we did not necessarily expect children's performance in Trial 2 to be identical to that in Trial 1. For example, it is possible that seeing the experimental procedure for a second time altered children's thinking about the task and thus caused them to respond differently. We therefore performed analysis focussed on differentiating whether specific basket choices (informed, uninformed, other) differed between Trials 1 and 2 i.e. does Trial (1 or 2) affect propensity to choose an informed, uninformed or other basket. We analysed this using a generalised linear mixed model (GLMM).

Our GLMM (family = binomial) had a dependent variable of choice correct (1 or 0 as in previous analyses). The fixed effects were trial number (sum coded with Trial 1 set to 1, Trial 2 to -1), age (in days, centred and scaled to measure thousands of days) and the

interaction of trial number and age. We included a random intercept for participant and a by-participant random slope for trial number. We found no main effect of age ($b = 0.166$, $SE = 0.341$, $Z = 0.488$, $p = .625$), trial number ($b = -0.201$, $SE = 0.477$, $Z = -0.421$, $p = .674$) or their interaction ($b = 0.026$, $SE = 0.273$, $Z = 0.094$, $p = .925$) on choice correct.

Verbal Responses

If SLSs based on explicitly metacognitive understanding were being employed, we would expect participants to choose the same basket as the informed demonstrator in addition to providing a verbal explanation as to why they made this choice. However, we acknowledge that an individual may use an understanding-based SLS without having the verbal ability to explain their understanding in which case there would be no verbal response to substantiate their basket choice (this is likely to be more common in children). Nevertheless, we analysed verbal responses to the question: “Why did you choose this basket?” which was asked immediately after children/adults chose one of the four baskets. Verbal responses were placed into seven categories (Table 5.7A – Trial 1, Table 5.7B – Trial 2) based on their content, as follows: one – informed (copying the informed demonstrator without mentioning their additional knowledge), two – explicitly metacognitive (copying the informed demonstrator with reference to their additional knowledge), three – uninformed (copying the uninformed demonstrator), four – shift (selecting one of the two baskets not chosen by either demonstrator), five – other, reasoned (a seemingly reasoned response but one in which the reasoning was incorrect, inappropriate or based on an incorrect/misremembered fact), six – unreasoned (a response was given but it was incorrect/nonsensical) and seven – no response (no verbal response was provided). For Categories 1 to 3 we also coded whether or not the verbal response corresponded to the basket choice made i.e. for Categories 1 (informed) and 2 (explicitly metacognitive) did participants choose the same basket as the informed demonstrator (the correct basket) and for Category 3 (uninformed) did they choose the same basket as the uninformed demonstrator (the incorrect basket). We deemed this correspondence to be an indicator that an individual was providing a true verbal justification for their basket choice, rather than producing a post-hoc justification for a choice they later determined should have been made.

Table 5.7

Number of Participants in Trials 1 (Table 5.7A) and 2 (Table 5.7B), Aged 3-8 and Adult, Who Made Each of Seven Verbal Response Types (1-7). Percentage of Participants Shown in Brackets. For Response Types 1-3 we Also Display Whether Participants Chose a Basket Which Aligned with Their Verbal Response

A.

Age (whole years, number of participants in brackets)	Trial 1									
	Verbal response types 1-7 – number of participants (% of participants in brackets)									
	1. Informed total	Informed & aligned (correct) basket choice	2. Metacognitive total	Metacognitive & aligned (correct) basket choice	3. Uninformed total	Uninformed & aligned (incorrect) basket choice	4. Shift total	5. Other, reasoned	6. Unreasoned response	7. No response
3 (20)	0	0	0	0	0	0	0	0	17 (85.0)	3 (15.0)
4 (21)	0	0	0	0	1 (4.8)	1 (4.8)	0	1 (4.8)	17 (81.0)	2 (9.5)
5 (20)	0	0	0	0	1 (5.0)	1 (5.0)	0	0	16 (80.0)	3 (15.0)
6 (19)	2 (10.5)	1 (5.3)	0	0	2 (10.5)	1 (5.3)	1 (5.3)	2 (10.5)	12 (63.2)	0
7 (20)	0	0	1 (5.0)	0	0	0	0	4 (20.0)	15 (75.0)	0
8 (19)	0	0	1 (5.3)	1 (5.3)	1 (5.3)	1 (5.3)	1 (5.3)	4 (21.1)	10 (52.6)	2 (10.5)
adult (18)	0	0	15 (83.3)	15 (83.3)	1 (5.6)	1 (5.6)	0	2 (11.1)	0	0

B.

Age (number of participants in brackets)	Trial 2									
	Verbal response types 1-7 – Number of participants (% of participants in brackets)									
	1. Informed total	Informed & aligned (correct) basket choice	2. Metacognitive total	Metacognitive & aligned (correct) basket choice	3. Uninformed total	Uninformed & aligned (incorrect) basket choice	4. Shift total	5. Other, reasoned	6. Unreasoned response	7. No response
3 (20)	2 (10.0)	2 (10.0)	0	0	0	0	1 (5.0)	1 (5.0)	13 (65.0)	3 (15.0)
4 (21)	0	0	0	0	1 (4.8)	0	0	0	18 (85.7)	2 (9.5)
5 (20)	0	0	0	0	0	0	0	3 (15.0)	16 (80.0)	1 (5.0)
6 (19)	1 (5.3)	1 (5.3)	0	0	0	0	0	3 (15.8)	15 (78.9)	0
7 (20)	1 (5.0)	0	2 (10.0)	2 (10.0)	0	0	1 (5.0)	5 (25.0)	11 (55.0)	0
8 (19)	0	0	2 (10.5)	2 (10.5)	3 (15.8)	3 (15.8)	0	3 (15.8)	10 (52.6)	1 (5.3)
adult (18)	0	0	18 (100.0)	18 (100.0)	0	0	0	0	0	0

We were especially interested in the proportion of participants who made the correct basket choice in addition to giving an “explicitly metacognitive” verbal response i.e. a response showing that they had used their understanding of the informed demonstrator’s mind to help them achieve task success. Although some individuals who chose the same basket as the uninformed demonstrator were able to explain the fact that they made this choice, we would not class these individuals as using an understanding-based SLS which utilised explicit metacognition. This demonstrator was not privy to any more knowledge than the participants themselves therefore if participants were utilising their understanding of the knowledge states of others to devise an explicitly metacognitive, understanding-based SLS they would not be expected to copy this demonstrator.

As expected, most adults provided an explicitly metacognitive verbal response and a corresponding correct basket choice (15 and 18 in Trials 1 and 2 respectively – 83% and 100% of all adult participants). No adults gave an explicitly metacognitive verbal response in the absence of a correct basket choice. In contrast, children’s verbal responses did not indicate anything more than occasional use of explicitly metacognitive, understanding-based SLSs at any of the age groups tested in either Trial 1 or 2. Only one child (aged 8) chose the correct basket and provided an explicitly metacognitive verbal response in Trial 2, and only four children (two aged 7, two aged 8) did so in Trial 2 (Tables 5.7A and 5.7B – “Metacognitive & aligned (correct) basket choice”). Furthermore, only one child (aged 7, Trial 1) provided an explicitly metacognitive verbal response but did not choose the correct basket. The predominant child response type was “unreasoned response” which was given by at least 52% of children in each age group (Trials 1 and 2).

Discussion

We aimed to determine whether there is a shift in social information use which coincides with developments in explicit metacognition through detecting whether children’s ability to copy an informed (rather than an uninformed) demonstrator changes during the ages of 3 to 8. Identification of such a shift would have been consistent with the idea that explicit metacognition may aid more successful information use through the derivation of understanding-based, explicitly metacognitive social learning strategies (Dunstone & Caldwell, 2018). Following the results of our previous study (Chapter 4), it was important that this follow-up allowed those with explicitly metacognitive understanding to achieve task success and thus perform as expected according to our paradigm. Establishing a task which adults would interpret as intended, and score highly on, was therefore an important goal of the current study.

Adult performance was as expected and, unlike child performance, was consistent

with the use of explicit metacognition to devise understanding-based SLSs. In contrast to the findings from our first study (Chapter 4), significantly more adults (than would be expected by chance) chose the same basket as the informed demonstrator: all but two adults (88.9%) were correct on the first trial, and all adults on the second trial. Furthermore, all but three adults (83.3%) in Trial 1, and all adults in Trial 2, provided a verbal explanation consistent with the use of an understanding-based, explicitly metacognitive social learning strategy in addition to choosing the correct basket.

In contrast, children of all ages chose the informed, uninformed and other baskets at chance levels which indicated random responding or responding based on misunderstanding/misinterpretation of the information presented. This result was somewhat supported by the finding that between 52.6 and 85.7 percent of children (across ages 3-8, Trials 1 and 2 – Tables 5.7A and 5.7B) gave an unreasoned verbal response when asked why they selected their chosen basket. Moreover, even children who did choose the correct basket were generally not able to support this action with a verbal justification indicative of explicit metacognitive understanding (10.5 percent or less across all age groups for Trials 1 and 2 – Tables 5.7A and 5.7B) and no child below age 7 provided such a justification. It should be noted that, although the verbal responses lend some support to the lack of correct responding in children, consideration of these unreasoned responses alone would not be sufficient evidence to determine that an understanding-based, explicitly metacognitive SLS was not used. It is possible that children could form and use such a strategy without having the ability to verbalise it. Additionally, we did not find any evidence that propensity to copy the informed demonstrator differed according to age in children: there was no effect of age on repeating of the informed demonstrator's choice (Trials 1 and 2) or on total score (Trials 1 and 2 combined). In conclusion, we did not find any evidence from the children's physical basket choices, or their verbal reporting to support the proposed shift in social learning strategy.

In our first study (Chapter 4), only 60% of adults actually chose the correct basket (which was not significantly above chance level) although most adults (83.3%) who chose the correct basket provided verbal responses consistent with explicitly metacognitive understanding. Yet in the current experiment adult performance was close to ceiling, and thus also distinct from that of the children, who did not copy the informed demonstrator at levels significantly different to chance. The difference in adult performance between the two experiments was likely the result of our making the correct response more transparent in the current study. Factors such as reducing the need to pit one social learning strategy against another (and therefore reducing the number of demonstrators) probably resulted in more adults interpreting the task as we had intended (additional reasons for the poor performance

of adults in our first study are discussed in Chapter 4). It therefore seems that we sufficiently addressed the issues identified with the previous task and that the paradigm used here enabled participants who understood the knowledge state of the informed demonstrator to selectively copy this demonstrator (as intended). Questions therefore remain regarding children's poor performance – why did children fail to use an understanding-based, explicitly metacognitive SLS? Children did not achieve this even by age 8, despite surpassing milestones in the development of capacities we had thought might be relevant (understanding of false belief, how evidence relates to knowledge and/or partial knowledge exposure). We suggest possible reasons for this in the subsequent paragraphs.

There is much literature showing that children can answer questions regarding an individual's knowledge based on that individual's visual access from age 3 or 4 (Pillow, 1989, Experiment 1; Pratt & Bryant, 1990, Experiments 2 and 3; Wimmer et al., 1988), and can also identify whether an individual with or without visual access to the contents of a box knows the contents or should be asked a question regarding the contents (Pillow, 1989, Experiment 2; Pratt & Bryant, 1990, Experiment 1; Schmidt & Pyers, 2011). In our task, we would therefore have expected most children to understand that the informed demonstrator, with visual access to the contents of the eggs, knew which basket contained the chicks. However, asking children to use this information, in preference to outwardly similar information from the uninformed demonstrator (i.e. looking inside each of the four baskets, but not the eggs themselves) has not been fully investigated and could present more of a challenge than understanding the informed demonstrator's knowledge state alone. Nevertheless, by age 5 or 6 children pass tasks which require a greater understanding of the subtleties required for individuals to gain different types of knowledge e.g. that an individual cannot necessarily know what image is depicted if they only have visual access to a small part of it (Chandler & Helm, 1984; Taylor, 1988). We would therefore have expected that these developing abilities around age 5 or 6 may aid children's detection of the differences between the informed and uninformed demonstrator's levels of visual access and therefore support them in concluding that only the informed demonstrator's access provided relevant information. Of course, we cannot isolate the contribution that different cognitive developments may, or may not, have made but our results did not provide any indication of differential task performance in line with any of the developments discussed.

The changes made in the current study likely lessened memory demands – primarily by decreasing the number of uninformed demonstrators from three to one. We also made some further, small edits to the task “story” presented to children in order to reduce the cognitive load. This included the removal of unnecessary details (and therefore related language) such as having a character (the fox) which needed to be paid in order to search

the locations. We also eliminated non-essential task materials such as the laminated board, plastic coins, and the chest which contained them (see Chapter 4). Memory ability may be important for devising SLSs in practise. However, reducing this element of the study still enabled us to address our main research question – whether a beneficial shift in social information use could be due (at least in part) to the development of human-unique explicitly metacognitive understanding.

Despite the aforementioned changes to memory demands, the current task did not provide any evidence that children were considering the information provided by the informed demonstrator as preferable at any age. Even with our changes, it remains possible that the executive function demands were too high to allow children to use explicitly metacognitive understanding to formulate an understanding-based SLS. This is something we could continue to investigate in future experiments. However, using the results from this study, it was possible to further examine the potential influence of memory by analysing the data from only children who passed all our memory questions (and therefore for which memory demands were not a barrier to task success). The findings were as in the full dataset – no difference in social information use with age, and chance responding for each basket type. This result was also in line with the findings from our previous experiment in which children who passed all the memory questions did not copy the informed demonstrator's basket at above chance levels. It therefore appears that memory demands alone were not preventing children from using the informed demonstrator's information.

It could be argued that lessening the executive function demands made the experiment less ecologically valid. We would agree with this critique because in “real-life” learning situations the ability to discount irrelevant information may be an important aspect of the capacity to derive understanding-based, explicitly metacognitive SLSs which are routed in accurate analysis of the available evidence. For example, an ability to identify that one individual has visual access to useful information (e.g. through understanding that they have knowledge due to their unique visual access), in the absence of an ability to filter out different sources of (potentially conflicting) information, would not enable a learner to devise the most appropriate understanding-based, explicitly metacognitive SLS. Our removal of numerous uninformed demonstrators, providing irrelevant information, may have therefore eliminated part of the difficulty in deriving an SLS specific to the current task situation. However, even with this change, we did not find that children chose to copy the informed demonstrator above chance levels or in conjunction with the increasing development of explicit metacognitive understanding. If we had found evidence to this effect, further research could then have addressed how additional cognitive abilities, such as advanced executive functions, may be required in order to generate SLSs which are functional outside

of the laboratory.

In Chapter 4 we discussed how being exposed to the potential contents of the eggs and baskets (i.e. one example egg with a chick inside, and one with a yolk inside) may have given children a “feeling of knowing” which basket contained the eggs with chicks inside. This would align with partial exposure searching tasks (Kloo et al., 2017; Rohwer et al., 2012; Sodian & Wimmer, 1987) in which children are reported to experience implicit feelings of knowledge if they have been pre-exposed to the item which is to be hidden. In this second study we again showed children the two example eggs at the start of the task and therefore it was possible that they may have had a feeling of knowing which of the four baskets contained the chicks. Yet, as stated in Chapter 4, if this phenomenon was solely responsible for children’s poor task performance, we would have expected to observe an improvement from age 6 (around the age children normally pass partial exposure tasks). As we did not observe such an improvement, it seems unlikely that any partial exposure element of the task was a major barrier to success.

Related to the above, performing successfully at the current task involves putting aside any preconceptions as to which basket might contain the sought-after chicks. One must be open to the possibility that they could be in any one of the four baskets and that either the informed or uninformed demonstrators may have useful information regarding this. Studies have shown that children find considering multiple possibilities difficult until at least age 5-6 (Beck et al., 2006) and still have difficulty at age 6 plus (Sophian & Somerville, 1988) if they need to consider multiple possibilities for an event which has already happened (e.g. if an object has already been hidden inside one of four cups and there are multiple possibilities as to which it could be inside). For example, even those aged 7-8 (Robinson et al., 2006) find this difficult when they additionally need to consider another individual’s thoughts in response to multiple possibilities. Although these studies differ from ours in that they involve the need to act according to an understanding that multiple possibilities need to be accounted for at once (e.g. that a mouse could run down one of two slides and a mat should be placed at the end of each, Beck et al., 2006), the underlying cognitive abilities required may be related to those needed to consider multiple possibilities in our task. Results have shown that considering multiple possibilities for an event which has already occurred may be more difficult because children can easily imagine a possible outcome (Beck et al., 2011) which may interfere with their ability to evaluate their own knowledge and all possible outcomes. This is similar to the feeling of knowing children may experience in the partial exposure tasks described above (Kloo et al., 2017; Rohwer et al., 2012; Sodian & Wimmer, 1987). However, considering multiple possibilities for an event which has already occurred (such as the situation presented in the current task) has also been postulated to

increase metacognitive demands because one needs to represent both what one knows and what one *could* know (Beran et al., 2012). Representing what could be known is a form of counterfactual thinking (“thinking that considers alternative possibilities for an event or behaviour in the past”, Nakamichi, 2019) and this is thought to continue to develop throughout middle childhood and even into adolescence (Rafetseder et al., 2010, 2013; Rafetseder & Perner, 2014). It may therefore be that a lack of developed counterfactual thinking affected children’s ability to consider the multiple possibilities in our task and thus may be implicated in the poor performance of children of children aged 3-8.

In each trial of this study, both the informed and uninformed demonstrator made actions to indicate that they were looking inside all four baskets (although the informed demonstrator additionally peeked inside an egg from each). The baskets contained the closed eggs therefore it should have been evident that simply looking inside these, without opening any eggs, would not lead to knowledge as to the contents of the eggs. We aimed to sit children at such a height that they could easily see into the baskets themselves and so would have full visual access to the closed eggs in each. Yet it is possible, although we suspect unlikely, that some children were unable to see inside the baskets, and we did not test what each participant could see. In such a scenario, children might have mistakenly thought that looking into the baskets (an action performed by both the informed and uninformed demonstrator) would lead to knowledge of the eggs’ contents (although children were shown that chicks or yolks were hidden *inside* eggs at the start of the game). In this case children may have considered the informed and uninformed demonstrator equal in regard to their knowledge of which basket contained eggs with chicks inside. Yet, we think that this is unlikely for a number of reasons. Firstly, the baskets selected by both the informed and uninformed demonstrators were not chosen by children at above chance levels. Furthermore, the informed demonstrator very clearly picked up an egg from each basket and held each up above the top of the basket whilst discreetly opening it to peek inside. There is no doubt that these actions were visible to all children and they should have been very obviously distinguishable from those of the uninformed demonstrator (i.e. simply looking inside the baskets).

The same four sticker colours were used for the yellow baskets in Trial 1 and the green baskets in Trial 2. This may have increased the possibility of children developing preferred or dispreferred colours in Trial 1 which could have then influenced their choice of basket in Trial 2. I aimed to minimise this by not revealing which basket from Trial 1 contained the chicks until both Trials 1 and 2 were complete. However, it is possible that this still had some influence on the results and future studies would need to eliminate this possibility.

Chapter 6: General Discussion

Summary of Thesis Aims and Hypotheses

The experimental work presented in this thesis has examined the apparent distinctiveness of human cumulative culture. I have discussed how this human proclivity to build on our achievements over generations appears to be in sharp contrast to the limited, and somewhat contentious, examples recently demonstrated in nonhumans (Jesmer et al., 2018; Sasaki & Biro, 2017; Schofield et al., 2017). However, scientists are yet to determine an explanation for the uniqueness of human cumulative culture, although numerous proposals have been made (see introduction pp.11-19). Human-unique cognitive capacities have been strongly implicated (e.g. Dean et al., 2012; Tennie et al., 2009), but no consensus reached.

The overarching aim of this thesis was therefore to use a novel developmental approach to investigate the human-unique cognitive capacities which may underlie human propensity for cumulative culture. We know that the capacity for cumulative culture presents at some point during development therefore performing experimental work with children offers unique advantages over work with adult populations (see introduction pages 26-33). The work I have presented rests on the principle that if children begin to use social information differently in line with the development of a particular human cognitive ability then this indicates that the ability may be a fundamental requirement for cumulative culture in the context presented. By “differently” I refer to using the information in a manner which would be considered advantageous/effective for information ratcheting. For example, as in Chapters 2 and 3, copying correct, and ignoring incorrect, task responses revealed by a demonstrator(s) (despite constraints, which differed depending on the condition/task presented) in order to outperform the demonstration(s). Or alternatively, as in Chapters 4 and 5, copying a demonstrator who is suitably informed as to the location of a reward in order to find the reward oneself.

Our aim to investigate cumulative culture across human development was aided by the use of a recently proposed method (Caldwell et al., 2020, see introduction pages 22-26) which eliminated many of the logistical challenges associated with experimentally investigating cumulative culture across a wide age range (as discussed in the introduction pages 19-26). Chapters 2 and 3 have presented the first documented studies which utilise this new method to investigate ratcheting potential – in this case to look at age-related changes, and thus the involvement of developing cognitive capacities. Furthermore, the adaptability of this method has enabled us to investigate the impact of task context and the challenges involved in using information from multiple social models. These factors are

important when considering the cognitive capacities required for ratcheting as well as the limitations on ratcheting in populations without human adult level cognitive abilities (e.g. children and nonhumans). In regards to multiple models (discussed further below), we have outlined how cognitive capacities may actually underlie the proposed relationship between increases in population size and cumulative culture (Dean et al., 2014; Henrich, 2004; Richerson, 2013).

In addition to exploring the cognitive requirements for cumulative culture in relation to task context and multiple models, we utilised the aforementioned developmental approach to test a recent hypothesis for the uniqueness of human cumulative culture – the explicitly metacognitive cumulative culture hypothesis (Dunstone & Caldwell, 2018, see also introduction pages 14-17). The studies presented in Chapters 4 and 5 represent the first experiments to test this hypothesis in children. Specifically, and in line with this hypothesis, I analysed whether social information is used more flexibly as human-unique capacities for explicit metarepresentation (in this case, theory of mind) develop. The proposed shift, from the use of default social learning strategies to understanding-based, explicitly metacognitive social learning strategies may enable cumulative culture through the selective copying of more knowledgeable others.

Summary and Discussion of the Work Presented in Chapters 2-5

Chapter 2

As stated above, recent research has suggested that some nonhumans may display cumulative culture (Jesmer et al., 2018; Sasaki & Biro, 2017; Schofield et al., 2017), albeit under restricted circumstances. This is especially interesting when considered in conjunction with experimental work (Caldwell & Millen, 2009; Zwirner & Thornton, 2015) which has shown that the cognitive mechanisms required for ratcheting may differ depending on the task context/behaviour being transmitted. We might therefore expect that a population with a less extensive suite of cognitive abilities and behavioural proclivities (e.g. nonhumans or young children) may show ratcheting in a more restricted range of circumstances than a population with a greater range of abilities (e.g. adults) who can access more information. In this respect, the capacity for cumulative culture would not be considered “all or nothing” (as it often seems to be portrayed in the literature). Instead, cumulative culture would be expected to manifest in a different range of circumstances in different populations. In Chapter 2, we investigated this in children aged 3 to 6.

We posited that because children’s cognitive abilities typically increase with age, a concomitant increase in the range of contexts within which they can demonstrate the potential for cumulative culture should be expected. Such a finding would support the idea

that cumulative culture depends on the interaction between cognitive ability and the learning context. Whilst previous experimental work in children has lent some support to this idea (see discussion of relevant literature in Chapter 2), we believe that we are the first to directly investigate age-related changes in children's potential for ratcheting. In order to achieve this we employed the aforementioned new method (proposed by Caldwell et al., 2020, see also Chapter 2 and introduction pages 22-26), which allowed us to measure the potential for cumulative culture within individual children whom we exposed to task solutions of multiple, different levels of success. We used children's performance to infer the theoretical outcome of a linear transmission chain, consisting of these different level solutions (corresponding to theoretical positions in the chain), and thus to determine children's potential for ratcheting. The task solutions presented to children were manipulated in two conditions which constituted different contexts (designed to emulate real-world social information use) and allowed us to determine the effect of both task context and age (a proxy for cognitive ability) on potential for ratcheting. In our enduring cues condition the task solutions remained present within the task environment for the participant to utilise to improve their performance, whereas in our transient cues condition these solutions existed for a limited period and so presented greater cognitive demands (particularly on memory).

We found support for the hypothesis that the potential for a population to show cumulative culture depends both on cognitive capacity and the context/behaviour in question. There was evidence of the potential for cumulative culture in all age groups (3-6) when cognitive demands on memory were low (enduring condition). However, under higher memory demands (transient condition) only children in the oldest age group (age 6) were able to improve their performance following the observation of higher-success task solutions, and thus demonstrate the potential for ratcheting. Older children, with more developed cognitive capacities in comparison to younger children, were therefore more equipped to access and use information in the more demanding, transient condition. These results also add support to the hypothesis that examples of cumulative culture may be found in many species, under restricted circumstances, but that cumulative culture is likely to be present across a broader range of contexts in humans due to the possession of enhanced cognitive abilities.

The cognitive ability restricting performance in younger children was memory, and most likely working memory (see our discussion of this in Chapter 2, pages 43-44 and 62-63). Although social (Boyd & Richerson, 1996; Dean et al., 2014; Tomasello et al., 1993, 1994) and physical (De Oliveira et al., 2019; Osiurak et al., 2016; Osiurak & Reynaud, 2020) cognition have been proposed to account for the uniqueness of human cumulative culture (see introduction pages 11-19), more general cognitive abilities such as working memory

have received little attention (although see Fay et al., 2019). This is likely to be due, in part, to a lack of firm evidence that humans have an enhanced working memory compared to nonhumans. However, the small amount of research available does suggest that humans can make better use of their raw working memory capacity than nonhumans through utilising various strategies such as metacognitive storage strategies (see Carruthers, 2013, and further discussion of this point in Chapter 2, pages 62-63). In our study, young children, who are unlikely to have developed strategies sufficient to support their memory on this task (Armitage et al., 2020; Bulley et al., 2020), did not show ratcheting in a condition in which their working memory was taxed and, as outlined above, we would likely find similar performance in nonhumans. However, older children, likely in possession of enhanced memory, were able to utilise the information more successfully and showed the potential for ratcheting from age 6.

Yet, although working memory ability appears to have affected ratcheting potential in this study, I do not advocate for working memory, or memory ability generally, being the sole cognitive factor separating human and nonhuman propensity for cumulative culture (as has been claimed for other cognitive abilities such as social learning mechanisms, see introduction pages 12-14). Rather, that in this particular task context, enhanced working memory in older children (but not younger) enabled ratcheting, and that this lends support to the proposal that humans may show cumulative culture in a greater number of contexts due to human-unique cognitive abilities. Nonetheless, the more cognitively challenging, transient condition was designed to emulate a situation which might be experienced in a real-life social learning scenario. Consequently, working memory may be a somewhat overlooked cognitive capacity which could be important for ratcheting in certain contexts. I return to this point later in the general discussion.

Chapter 3

In Chapter 3 we investigated whether human-unique cognitive abilities may be responsible for the positive relationship researchers have identified between population size and cumulative culture (see introduction pages 18-19). That is, the results of mathematical models (Henrich, 2004; Kobayashi & Aoki, 2012; Powell et al., 2009; Shennan, 2001) and ethnological studies (Collard et al., 2013; Kline & Boyd, 2010), recently backed-up by experimental work (Derech et al., 2013; Kempe & Mesoudi, 2014; Muthukrishna et al., 2014), which have suggested that the uniqueness of human cumulative culture may result from the formation of large, interconnected populations. Two underlying mechanisms have been proposed to enable this information from larger populations to be exploited for cumulative culture: selectively learning from the best of a larger group of individuals (Derech et al., 2013;

Henrich, 2004), and (more recently) combining information from multiple models (Kempe & Mesoudi, 2014; Muthukrishna et al., 2014). However, the existence of these mechanisms does not discount a cognitive explanation for uniquely human cumulative culture because such mechanisms may require adult-level cognitive capacities. This does not appear to have been considered by most researchers studying effects of population size, although two studies in adults do highlight that using information from multiple models may be cognitively demanding (Caldwell & Millen, 2010; Fay et al., 2019).

As in Chapter 2, we utilised a child population and the aforementioned method (Caldwell et al., 2020). However, we instead set out to determine whether ratcheting via integrating information from multiple models is cognitively challenging and thus may require the development of human-unique cognitive capacities. We posited that utilising information from more than one individual in real life likely requires bringing together information which is separated in time and space (Subiaul et al., 2015). Thus we expected that capitalising on information from multiple models would draw upon cognitive skills such as spatial translation, memory and evaluation of alternative information, increasing cognitive load. It followed that populations with more limited domain-general cognitive processing abilities and a lack of compensatory strategies, such as children and possibly nonhumans, would be likely to find use of this information challenging (Carruthers, 2013).

We expected that children's ability to utilise information from multiple models effectively would increase with age and that such age-related changes would enable us to further understand whether human-unique cognitive processes were required in order to ratchet the information presented. To investigate this, we measured use of information provided by two models in children aged 5 to 10; the information was separated in time and/or space over four conditions presenting different constraints – Temporal, Spatial and Temporal-Spatial displacement (plus a no-displacement control). Furthermore, in order to measure the potential for cumulative culture we analysed whether children were able to outperform the highest single scoring model in the two-model population – effectively, creating a new trait variant – in each of the four conditions.

In line with our predictions, we found that children's performance (using our various measures of information use) increased with age in the three displacement conditions but not in the no-displacement control in which performance was high across the entire age range. Furthermore, children showed the potential for cumulative culture across a greater number of the displacement conditions with increasing age – children aged 7-10 in all three displacement conditions, aged 6 in all but the Temporal-Spatial condition and those aged 5 did not show this potential in any of the displacement conditions. This study was therefore unique in demonstrating that the types of cognitive constraint presented, when learners are

faced with information from multiple models, may differentially affect ability to make efficient use of the information at different ages i.e. in our study younger children (with less developed cognitive abilities) found utilising the displaced information more difficult. Furthermore, the study was the first to show that the ability to improve on the highest single model score within a multiple model population, showing the potential for cumulative culture, increased with age and was dependent on the level of cognitive load presented (the task constraints). These results suggest that humans are able to exploit information from multiple models and drive cumulative culture due to enhanced cognitive abilities. Therefore, a cognitive explanation may underlie the population-size effect.

As in the work presented in Chapter 2, we found that both age (i.e. stage of cognitive development) and task context affected children's potential for ratcheting. Considering the two studies together, it seems that these two factors may act to determine when ratcheting will occur. Our findings from this second study therefore further support our statement that cumulative culture may occur in a broader range of circumstances in humans due to the possession of enhanced cognitive abilities. Furthermore, this latter study highlights the cognitive demands associated with utilising information from multiple models under ecologically valid contexts (e.g. spatial and temporal separation). Capitalising on information from multiple models could therefore be one of a potentially larger number of situations from which humans have been uniquely able to extract and build upon information which would be inaccessible to those with less enhanced cognitive capacities. However, even in the case of adults we obviously expect there to be constraints on accessing and using information (as found in Caldwell & Millen 2010; Fay et al., 2019) which limit our potential for cumulative culture in the absence of additional strategies.

In our previous analysis of Chapter 2, I have discussed the potential involvement of working memory, and human-unique strategies to enhance this, both for use of information in the task context presented and for cumulative culture more broadly. Working memory is also relevant to discussion of the work described here because the temporal, spatial and temporal-spatial displacement of the information presented from the two models increased the cognitive load. For example, accurately using information from the temporal displacement conditions required the holding and manipulation of information in mind. Moreover, the spatial conditions will have drawn upon mentally taxing skills such as mental translation (Frick et al., 2013; Levine et al., 1999; Shepard & Metzler, 1971) and an understanding of dual representation (DeLoache, 1989, 1991, 2000) in order to relay information between the models' search spaces and the child's. Developing working memory (Best & Miller, 2010; Diamond, 2013; Garon et al., 2008) and, moreover, human-unique methods of enhancing this such as metacognitive storage strategies (e.g. mnemonic

devices, Jurowski et al., 2015, or inner speech, Carruthers, 2013; Cowan, 2008) may thus have contributed to the higher performance found in older children and prevented young children's effective use of the information. As stated in regard to Chapter 2, this does not mean that human-like working memory is an essential requirement for cumulative culture. However, it does suggest that this may offer an advantage when utilising information from various different sources, such as multiple models, and act to enable cumulative culture in such circumstances.

Metacognition is a further cognitive development which we expect might have supported the improved use of information from multiple different sources in our displacement conditions from age 6. As mentioned above, metacognitive processes can aid memory, but they may also allow for greater monitoring of the task and control over one's behaviour (Bryce & Whitebread, 2012). Moreover, a growing understanding that partial perceptual information cannot lead to full knowledge formation from age 6 might have allowed children greater understanding of their own lack of knowledge (Kloo et al., 2017; Rohwer et al., 2012; Sodian & Wimmer, 1987) and the value of information from each model, enabling more flexible use of this information. Although some nonhuman species have been found able to monitor their own uncertainty, and make simple responses accordingly (Beran et al., 2006; Bohn et al., 2017; Call, 2010; Call & Carpenter, 2001; Couchman et al., 2010; Hampton et al., 2004), it is unlikely that the metacognitive ability measured in these tasks would enable the devising and using of metacognitive strategies suitable for improving task performance in our study (for further discussion of this see Chapter 3, pages 97-102). The age-related findings in our task therefore highlight the potential role of human-unique metacognitive abilities for cumulative culture through aiding memory, task monitoring and control, and more selective information use.

In relation to the final point above, explicit metacognition (i.e. that involving higher-level, System 2, cognitive processes) has been proposed to enable human cumulative culture through the use of strategies which allow for more selective use of social information (Dunstone & Caldwell, 2018; Heyes, 2016, 2018, see also introduction pages 14-17). However, as we did not set out to investigate explicit metacognition in this task (nor did we measure it directly), and we only studied a set number of contexts, much further research is needed in order to understand the full involvement of this cognitive capacity. I further considered explicit metacognition in Chapters 4 and 5 and return to discuss it in the concluding section of this general discussion.

Chapters 4 and 5

In contrast to the preceding two chapters, in Chapters 4 and 5 I turned to the specific

investigation of social cognition, elements of which have been deemed responsible for the human propensity for cumulative culture since the early 90s (Tomasello et al., 1993, 1994; Tomasello, 1999). However, unlike this earlier work, we were interested in whether the human-unique ability to explicitly represent the content of others' minds (i.e. theory of mind and related human-unique developments) facilitates cumulative culture through enabling more strategic social information use.

It has recently been proposed that the capacity for explicit metacognition facilitates ratcheting by enabling humans to use social information more selectively than other animals (Dunstone & Caldwell, 2018; Heyes, 2016, 2018). Dunstone and Caldwell (2018) posit that this capacity for explicit metarepresentation (both mentalising and introspection) allows more selective use of information through altering the behaviour of social information receivers and senders. For example, an explicit understanding of another individual's knowledge state (an example of mentalising) could allow one to devise strategies to selectively copy individuals whom one knows to be the most knowledgeable (Dunstone & Caldwell, 2018). This is not expected to be possible for nonhumans. Although the social learning conducted by nonhumans is known to conform to certain adaptive rules (social learning strategies, or SLSs) – see Chapter 4, pages 103-110), these are unlikely to be explicitly represented or under voluntary control and thus would not allow for the most up-to-date information to be precisely copied (Heyes, 2016). In Chapters 4 and 5 we set out to test the hypothesis that explicit metacognition (in this case the ability to understand others' minds) may underlie cumulative culture through identifying whether development of this human-unique capacity coincides with a shift in the use of social information. We aimed to identify the stage in development at which such a shift occurs.

In Chapter 4, we looked for a shift from use of the type of social learning strategy we might expect to see in nonhumans (or young children) to one which required human adult-like explicitly metacognitive understanding (an understanding-based, explicitly metacognitive social learning strategy). Following our findings, in Chapter 5 we made some changes to the task design and instead focussed solely on identifying the stage in development at which children were able to utilise an understanding-based, explicitly metacognitive SLS. The tasks performed in both chapters relied upon the assumption that if a participant had an explicit understanding that one of a number of social demonstrators held additional task-relevant knowledge they would selectively copy this demonstrator. This would then lead to the participant achieving task success. We aimed to identify whether such a change in social information use aligned with the development of theory of mind and related capacities – understanding of false belief, how evidence relates to knowledge and/or partial knowledge exposure. We tested this in children aged 3 to 8, and adults.

The task presented in Chapter 4 required participants to watch one informed demonstrator and three uninformed demonstrators searching two locations for hidden rewards prior to searching themselves. Only one location was rewarded. Participants who understood that the informed demonstrator was privy to additional knowledge (due to visual access unavailable to the participant and the other demonstrators) were expected to selectively copy this demonstrator by choosing the same reward location i.e. to employ an understanding-based, explicitly metacognitive SLS. Those who did not have this understanding were instead expected to employ a common heuristic and “copy-the-majority” (Burdett et al., 2016; Haun & Tomasello, 2011; Van Leeuwen et al., 2018; Wilks et al., 2015) of uninformed demonstrators. We did not find any evidence for a shift from the heuristic strategy to an understanding-based, explicitly metacognitive strategy with increasing age in children – in no age group did children choose the correct location at above chance levels or give verbal responses indicative of explicitly metacognitive understanding.

However, the study in Chapter 4 was the first to investigate these questions in children and was therefore fairly exploratory – we did not know that participants with explicitly metacognitive understanding would selectively copy the informed demonstrator in this setup (although we predicted that they would). Actually, it appeared that on the whole they did not; adults, whom we expect to have explicitly metacognitive understanding, did not copy the informed demonstrator at above chance levels. This low performance seems to be due to adults not interpreting the task as intended because most of the adults who gave an explicitly metacognitive verbal response, and therefore presumably did interpret the task as we intended, chose the correct basket. As we had set out to determine at what stage children begin to use social information in an “adult-like” fashion, the poor adult performance meant that there was little to be interpreted from the child results. It thus appeared that our task was not suitable for detecting any proposed shift to adult-like social learning, hence we subsequently adapted it (Chapter 5) with the aim of devising a task in which explicitly metacognitive understanding would allow the correct rule to be applied and adults to perform as expected.

In Chapter 5 the task instead consisted of four search locations (one rewarded) and two demonstrators, one of whom was informed as to the location of hidden rewards and the other who was uninformed. Again, we expected participants with an understanding of the informed demonstrator’s additional knowledge to employ an understanding-based, explicitly metacognitive SLS and copy the informed demonstrator’s selected location. However, we also expected that it would be more obvious that this strategy would lead to task success because we did not pit it against a strategy we would expect to dominate in the absence of explicitly metacognitive understanding. Unlike in the study from Chapter 4, adults did

perform as would be expected if they were using a strategy based on explicitly metacognitive understanding – most adults chose the same basket as the informed demonstrator in addition to providing a corresponding verbal response. It thus seems that this study was more suitable for detecting the proposed shift in social learning. Yet, in children there was still no evidence of a shift to use of an explicitly metacognitive strategy (i.e. choosing the same location as the informed demonstrator and/or providing verbal responses which indicated such understanding) with increasing age. Furthermore, the different reward locations (that selected by the informed demonstrator, the uninformed demonstrator or one of the remaining two locations) were selected at chance level in all age groups, indicating that participants were responding randomly or possibly according to a range of different, incorrect strategies.

Considering the results from both Chapters 4 and 5, we found no evidence to support the proposal that the explicit understanding of others' minds enables the aforementioned shift in social learning strategy (and thus more flexible use of social information) during the ages of 3 to 8. In Chapter 4 this was likely to be a result of issues with task interpretation because even adults did not perform as expected. However, following simplification of the study, in our second study (Chapter 5) there remained no evidence for such a shift. To eliminate the possibility that this finding remains due to issues with how the task is interpreted further work would be required. It is possible that even if children understood the knowledge state of the informed demonstrator, realising that this demonstrator should be copied may require further task-relevant experience which was not necessary for adults. It is also possible that children did not copy the informed demonstrator due to lack of a cognitive capacity/capacities which we have failed to consider.

Following our findings regarding the role of working memory in Chapters 2 and 3, in both Chapters 4 and 5 we examined the potential impact of ability to remember the selections made by the informed/uninformed demonstrators through performing additional analysis on a subset of the study population who answered all our memory questions correctly. The results from this sample did not reveal differential performance indicative of memory difficulties being responsible for children's poor performance.

Limitations and Future Directions

Here I focus on limitations of the experimental work presented in this thesis as a whole; limitations specific to the individual studies have been outlined in the chapter summaries above and/or in the relevant experimental chapters (2-5). Although this thesis has aimed to investigate the cognitive capacities which may underlie cumulative culture, none of the aforementioned studies have included standardised tests of the capacities I have

considered. Whilst this wasn't necessary in order to address our research questions, testing short-term and working memory (Best & Miller, 2010; Cowan, 2008; Diamond, 2013; Garon et al., 2008) could add support to the findings regarding the relationship between cognitive capacity and task context in Chapters 2 and 3, and moreover, the difficulties associated with utilising information from multiple models in Chapter 3. Working memory is recognised as one of the core executive functions - higher order cognitive processes which act in a top down manner to control and regulate behaviour, enabling this to become goal directed (Diamond, 2013; Garon et al., 2008). However, it can be difficult to discern whether a task utilises working memory or inhibition (the ability to avoid making a pre-potent response or performing a pre-potent behaviour) therefore it would also be informative to include a separate test of inhibition, which can be differentiated from working memory in the preschool and early primary school years (Roebbers, 2017). Additionally, it would be useful to include tests of social cognition, such as theory of mind (Beaudoin et al., 2020), alongside adapted versions of the studies conducted in Chapters 4 and 5.

Whilst performing such standardised tests could provide informative correlations between individuals (or groups) with certain capacities, and task performance (e.g. ratcheting in a particular task condition expected to require this capacity), this approach would still not definitively confirm that the capacity in question was a requirement rather than an additionally utilised capacity. Nevertheless, confirming that a particular ecologically valid task context taxes certain cognitive abilities could inform the design of future studies to further investigate the extent to which the capacity is implicated in human cumulative culture. Future studies could then measure children's potential for cumulative culture when such a capacity is blocked from use. As discussed in Chapters 2 and 3, it is likely that the exact cognitive capacities required for cumulative culture differ depending on the particular context (e.g. whether information remains easily assessable or requires storage in memory). However, there may be some contexts (e.g. utilising information from multiple models – Chapter 3) in which possessing certain cognitive abilities offers advantageous access to, and ability to improve upon, social information.

All the studies presented in this thesis use experimental methods to examine elements of cumulative culture – a group level process. Whilst we designed these studies to ensure as high a level of ecological validity as possible, no experimental study can fully replicate real-life social learning scenarios whilst providing interpretable results to address specific research questions. Therefore, we cannot possibly fully emulate elements such as population size and structure, or the timescales involved in cumulative culture, which can span many human generations. Our tasks are thus simplified (and necessarily abstract) models, and one could argue that any findings are context specific and not applicable to

contexts and populations beyond the laboratory. Going forward, it therefore remains important to substantiate any findings with long-term field studies of real populations.

Final Comments

The results from experimental Chapters 2 and 3 further highlight the need to consider context when aiming to understand the aspects of cognition which may aid ratcheting in different species. In humans, cumulative culture presents in a wide range of contexts and the cognitive mechanisms necessary for ratcheting in one situation/for one behaviour may not be so useful for another. That said, the studies in Chapters 2 and 3 were specifically designed to emulate contexts which are likely to frequently occur in real-life social learning scenarios (for example, having access to social information for a limited period or needing to combine multiple pieces of useful information from different sources), therefore any capacities strongly implicated could underlie many examples of human cumulative culture. For the information presented to participants in these chapters, we identified that memory (particularly, working memory) was a capacity likely limiting ratcheting performance in younger children (as discussed on pages 168 and 170-171). This may therefore be an important general cognitive capacity which has been largely overlooked in favour of physical (Osiurak et al., 2016) or social (Tomasello, 1999; Tomasello et al., 1993, 1994) cognition, and one which warrants further investigation. Working memory has previously been theoretically linked to cumulative culture (Balter, 2010; Coolidge & Wynn, 2001; Wynn & Coolidge, 2011). Yet only more recently has it been related to cumulative culture experimentally (Fay et al., 2019); specifically, it has been proposed that working memory limitations may restrict selective copying of the best trait variants (success-biased copying) as population size increases. Taking an alternative approach, the work presented in this thesis is the first to suggest that age-related developments in working memory (amongst other cognitive capacities) may enable ratcheting across a greater range of contexts in humans than nonhumans (including utilisation of information from multiple models within large populations).

The findings from our investigations into the relationship between explicit metacognition and cumulative culture are less tangible. Our work regarding children's potential use of understanding-based, explicitly metacognitive social learning strategies (Chapters 4 and 5) did not provide any evidence that explicit metacognition aided selective use of social information in children aged 3-8 (as has been proposed for adults in the explicitly metacognitive cumulative culture hypothesis, Dunstone & Caldwell, 2018), although there was evidence for this in adults (most adults who were able to articulate the rule also made the "correct" choice). The poor performance of children, and adults in Chapter 4 only,

could have been due to issues with the task(s) (as outlined above). However, this may also be because we specifically designed our tasks to determine the potential role of one, slightly controversial, aspect of explicit metacognition – mindreading or understanding of others’ minds (Dunstone & Caldwell, 2018; Heyes, 2016) i.e. asking “who knows?” Children’s poor performance across the 3-8-year age range suggested that the capacity to use understanding-based, explicitly metacognitive SLSs is not linked to the development of such perspective-taking capacities, at least in this task, and that additional abilities were required. It has also been proposed (in the explicitly metacognitive cumulative culture hypothesis, Dunstone & Caldwell, 2018) that introspection (understanding one’s own mind) may be highly relevant for strategic social learning, and thus cumulative culture, but our tasks were not designed to test this theory.

Dunstone and Caldwell (2018) posit that being aware of one’s own knowledge state, and when updating this knowledge could be advantageous, might allow for more strategic social information seeking (or a focus on innovation if the social information available is insufficient). In this respect explicit metacognition could be considered to assist in strategically directing learning according to information gained through reasoning in a specific situation. This would not necessarily need to involve social information but would be expected to enable habitual responses (which are commonly encountered in social learning situations) to be overcome. In other words, explicit metacognition of this sort may enable one to utilise information in situations which place high demands on executive functions, such as many of those involved in cumulative culture (Dunstone & Caldwell, 2018). Our findings from Chapter 3 speak to this hypothesis, although we did not set out to test it directly. We found that children were better at utilising information from multiple models (a situation presenting a high cognitive load) from around age 6, an age which coincides with significant advances in children’s understanding of their own knowledge (Kloo et al., 2017; Rohwer et al., 2012; Sodian & Wimmer, 1987). A growing understanding of the relationship between one’s own knowledge, the partial information provided in the task, and that needed to obtain the highest score may have encouraged more strategic use of information from each model. Explicit metacognition may have also aided memory, monitoring and control processes in the tasks presented in Chapters 2 and 3.

The experimental work presented in this thesis (Chapters 2-5) has explored cognitive capacities (both general and social) potentially responsible for the uniqueness of human cumulative culture. To this end, I have utilised innovative methods (e.g. Caldwell et al., 2020) which have enabled the study of ratcheting and social information use in children across a wide developmental range (ages 3 to 10). The work presented has not definitively determined the cognitive capacities underlying human cumulative culture. Yet, through

utilising a novel developmental approach, the thesis has shed valuable new light on cognitive mechanisms implicated in distinctively human culture.

References

- Allen, J., Weinrich, M., Hoppitt, W., & Rendell, L. (2013). Network-based diffusion analysis reveals cultural transmission of lobtail feeding in humpback whales. *Science*, *340*(6131), 485–488. <https://doi.org/10.1126/science.1231976>
- Aplin, L. M., Farine, D. R., Morand-Ferron, J., Cockburn, A., Thornton, A., & Sheldon, B. C. (2015). Experimentally induced innovations lead to persistent culture via conformity in wild birds. *Nature*, *518*(7540), 538–541. <https://doi.org/10.1038/nature13998>
- Armitage, K. L., Bulley, A., & Redshaw, J. (2020). Developmental origins of cognitive offloading. *Proceedings. Biological Sciences*, *287*(1928), 20192927. <https://doi.org/10.1098/rspb.2019.2927>
- Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men; research in human relations* (pp. 177–190). Carnegie Press.
- Atkinson, M., Renner, E., Thompson, B., Mackintosh, G., Xie, D., Su, Y., & Caldwell, C.A. (2020). Robust, source-independent biases in children's use of socially and individually acquired information. *Journal of Experimental Psychology: General*. <https://doi.org/10.1037/xge0000959>
- Balter, M. (2010). Did working memory spark creative culture? *Science*, *328*(5975), 160–163. <https://doi.org/10.1126/science.328.5975.160>
- Barr, D., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 1–43. <https://doi.org/10.1016/j.jml.2012.11.001>
- Beaudoin, C., Leblanc, É., Gagner, C., & Beauchamp, M. H. (2020). Systematic Review and Inventory of Theory of Mind Measures for Young Children. *Frontiers in psychology*, *10*, 2905. <https://doi.org/10.3389/fpsyg.2019.02905>
- Beck, S. R., McColgan, K. L. T., Robinson, E. J., & Rowley, M. G. (2011). Imagining what might be: Why children underestimate uncertainty. *Journal of Experimental Child Psychology*, *110*(4), 603–610. <https://doi.org/10.1016/j.jecp.2011.06.010>
- Beck, S. R., Robinson, E. J., Carroll, D. J., & Apperly, I. A. (2006). Children's thinking about counterfactuals and future hypotheticals as possibilities. *Child Development*, *77*(2), 413–426. <https://doi.org/10.1111/j.1467-8624.2006.00879.x>
- Beran, M. J., Decker, S., Schwartz, A., & Smith, J. D. (2012). Uncertainty Monitoring by Young Children in a Computerized Task. *Scientifica*, *2012*, 1–6. <https://doi.org/10.6064/2012/692890>
- Beran, M. J., Smith, J. D., Redford, J. S., & Washburn, D. A. (2006). Rhesus macaques

- (Macaca mulatta) monitor uncertainty during numerosity judgments. *Journal of Experimental Psychology. Animal Behavior Processes*, 32(2), 111–119.
<https://doi.org/10.1037/0097-7403.32.2.111>
- Best, J., & Miller, P. (2010). A Developmental Perspective on Executive Function. *Child Development*, 81(6), 1641–1660. <https://doi.org/10.1111/j.1467-8624.2010.01499.x>
- Blakey, K.H., Rafetseder, E., Atkinson, M., Renner, E., Cowan-Forsythe, F., Sati, S.J., & Caldwell, C.A. (2020). Development of Strategic Social Information Seeking: Implications for Cumulative Culture. *PsyArXiv*. <https://doi.org/10.31234/osf.io/4envx>
- Bohn, M., Allritz, M., Call, J., & Völter, C. J. (2017). Information seeking about tool properties in great apes. *Scientific Reports*, 7(1), 10923. <https://doi.org/10.1038/s41598-017-11400-z>
- Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. University of Chicago Press.
- Boyd, R., & Richerson, P. J. (1996). Why culture is common, but cultural evolution is rare. *Proceedings of the British Academy*, 88, 77–93. <https://doi.org/citeulike-article-id:1339814>
- Bryce, D., & Whitebread, D. (2012). The development of metacognitive skills: Evidence from observational analysis of young children's behavior during problem-solving. *Metacognition and Learning*, 7(3), 197–217. <https://doi.org/10.1007/s11409-012-9091-2>
- Buchsbaum, D., Gopnik, A., Griffiths, T. L., & Shafto, P. (2011). Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence. *Cognition*, 120(3), 331–340. <https://doi.org/10.1016/j.cognition.2010.12.001>
- Bulley, A., McCarthy, T., Gilbert, S. J., Suddendorf, T., & Redshaw, J. (2020). Children Devise and Selectively Use Tools to Offload Cognition. *Current Biology*, 1–8. <https://doi.org/10.1016/j.cub.2020.06.035>
- Burdett, E. R. R., Lucas, A. J., Buchsbaum, D., McGuigan, N., Wood, L. A., & Whiten, A. (2016). Do Children Copy an Expert or a Majority? Examining Selective Learning in Instrumental and Normative Contexts. *Plos One*, 11(10), e0164698. <https://doi.org/10.1371/journal.pone.0164698>
- Caldwell, C. A. (2018). Using Experimental Research Designs to Explore the Scope of Cumulative Culture in Humans and Other Animals. *Topics in Cognitive Science*, 1–17. <https://doi.org/10.1111/tops.12391>
- Caldwell, C. A., Atkinson, M., Blakey, K. H., Dunstone, J., Kean, D., Mackintosh, G., Renner, E., & Wilks, C. E. H. (2020). Experimental assessment of capacities for cumulative culture: Review and evaluation of methods. *Wiley Interdisciplinary Reviews: Cognitive Science*, 11(1), 1–15. <https://doi.org/10.1002/wcs.1516>

- Caldwell, C. A., & Millen, A. E. (2008a). Experimental models for testing hypotheses about cumulative cultural evolution. *Evolution and Human Behavior*, 29(3), 165–171. <https://doi.org/10.1016/j.evolhumbehav.2007.12.001>
- Caldwell, C. A., & Millen, A. E. (2008b). Studying cumulative cultural evolution in the laboratory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1509), 3529–3539. <https://doi.org/10.1098/rstb.2008.0133>
- Caldwell, C. A., & Millen, A. E. (2009). Social learning mechanisms and cumulative cultural evolution: Is imitation necessary? *Psychological Science*, 20(12), 1478–1483. <https://doi.org/10.1111/j.1467-9280.2009.02469.x>
- Caldwell, C. A., Renner, E., & Atkinson, M. (2018). Human Teaching and Cumulative Cultural Evolution. *Review of Philosophy and Psychology*, 9(4), 751–770. <https://doi.org/10.1007/s13164-017-0346-3>
- Caldwell, C. A., & Millen, A. E. (2010). Human cumulative culture in the laboratory: Effects of (micro) population size. *Learning & Behavior: A Psychonomic Society Publication*, 38(3), 310–318. <https://doi.org/10.3758/LB.38.3.310>
- Call, J., & Tomasello, M. (1999). A nonverbal false belief task: the performance of children and great apes. *Child Development*, 70(2), 381–395. <https://doi.org/10.1111/1467-8624.00028>
- Call, J. (2010). Do apes know that they could be wrong? *Animal Cognition*, 13(5), 689–700. <https://doi.org/10.1007/s10071-010-0317-x>
- Call, J., & Carpenter, M. (2001). Do apes and children know what they have seen? *Animal Cognition*, 3(4), 207–220. <https://doi.org/10.1007/s100710100078>
- Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, 12(5), 187–192. <https://doi.org/10.1016/j.tics.2008.02.010>
- Carruthers, P. (2008). Meta-cognition in animals: A skeptical look. *Mind and Language*, 23(1), 58–89. <https://doi.org/10.1111/j.1468-0017.2007.00329.x>
- Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, 32(2), 121–138. <https://doi.org/10.1017/S0140525X09000545>
- Carruthers, P. (2013). Evolution of working memory. *Proceedings of the National Academy of Sciences of the United States of America*, 110(SUPPL2), 10371–10378. <https://doi.org/10.1073/pnas.1301195110>
- Chandler, M. J., & Helm, D. (1984). Developmental Changes in the Contribution of Shared Experience to Social Role-Taking Competence. *International Journal of Behavioral Development*, 7(2), 145–156. <https://doi.org/10.1177/016502548400700203>

- Clements, W. A., & Perner, J. (1994). Understanding of Belief. *Cognitive Development*, 9(4), 377–395. <http://www.sciencedirect.com/science/article/pii/0885201494900124>
- Collard, M., Ruttle, A., Buchanan, B., & O'Brien, M. J. (2013). Population Size and Cultural Evolution in Nonindustrial Food-Producing Societies. *PLoS ONE*, 8(9), 1–6. <https://doi.org/10.1371/journal.pone.0072628>
- Collard, M., Vaesen, K., Cosgrove, R., & Roebroeks, W. (2016). The empirical case against the 'demographic turn' in palaeolithic archaeology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1698). <https://doi.org/10.1098/rstb.2015.0242>
- Coolidge, F. L., & Wynn, T. (2001). Executive functions of the frontal lobes and the evolutionary ascendancy of Homo sapiens. *Cambridge Archaeological Journal*, 11(2), 255–260. <https://doi.org/10.1017/S0959774301000142>
- Couchman, J. J., Coutinho, M. V. C., Beran, M. J., & Smith, J. D. (2010). Beyond stimulus cues and reinforcement signals: A new approach to animal metacognition. *Journal of Comparative Psychology*, 124(4), 356–368. <https://doi.org/10.1037/a0020129>
- Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? Nelson. *NIH Public Access*, 6123(07), 323–338. [https://doi.org/10.1016/S0079-6123\(07\)00020-9](https://doi.org/10.1016/S0079-6123(07)00020-9)
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13(4), 148–153. <https://doi.org/10.1016/j.tics.2009.01.005>
- De Oliveira, E., Reynaud, E., & Osiurak, F. (2019). Roles of Technical Reasoning, Theory of Mind, Creativity, and Fluid Cognition in Cumulative Technological Culture. *Human Nature*, 326–340. <https://doi.org/10.1007/s12110-019-09349-1>
- Dean, L. G., Kendal, R. L., Schapiro, S. J., Thierry, B., & Laland, K. N. (2012). Identification of the social and cognitive processes underlying human cumulative culture. *Science*, 335(6072), 1114–1118. <https://doi.org/10.1126/science.1213969>
- Dean, L. G., Vale, G. L., Laland, K. N., Flynn, E., & Kendal, R. L. (2014). Human cumulative culture: A comparative perspective. *Biological Reviews*, 89(2), 284–301. <https://doi.org/10.1111/brv.12053>
- DeLoache, J. S. (1989). Young children's understanding of the correspondence between a scale model and a larger space. *Cognitive Development*, 4(2), 121–139. [https://doi.org/10.1016/0885-2014\(89\)90012-9](https://doi.org/10.1016/0885-2014(89)90012-9)
- DeLoache, J. S. (1991). Symbolic Functioning in Very Young Children: Understanding of Pictures and Models. *Child Development*, 62(4), 736–752.
- DeLoache, J. S. (2000). Dual representation and young children's use of scale models. *Child Development*, 71(2), 329–338. <https://doi.org/10.1111/1467-8624.00148>
- Derex, M., Beugin, M. P., Godelle, B., & Raymond, M. (2013). Experimental evidence for the

- influence of group size on cultural complexity. *Nature*, 503(7476), 389–391.
<https://doi.org/10.1038/nature12774>
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1), 135–168.
<https://doi.org/10.1146/annurev-psych-113011-143750>
- Doherty, M. J. (2009). *Theory of mind : how children understand others' thoughts and feelings*. Hove : Psychology Press.
- Dunstone, J., & Caldwell, C. A. (2018). Cumulative culture and explicit metacognition: a review of theories, evidence and key predictions. *Palgrave Communications*, 4(1), 1–11. <https://doi.org/10.1057/s41599-018-0200-y>
- Enquist, M., Ghirlanda, S., Jarrick, A., & Wachtmeister, C. A. (2008). Why does human culture increase exponentially? *Theoretical Population Biology*, 74(1), 46–55.
<https://doi.org/10.1016/j.tpb.2008.04.007>
- Enquist, M., Ghirlanda, S., & Eriksson, K. (2011). Modelling the evolution and diversity of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1563), 412–423. <https://doi.org/10.1098/rstb.2010.0132>
- Enquist, M., Strimling, P., Eriksson, K., Laland, K., & Sjostrand, J. (2010). One cultural parent makes no culture. *Animal Behaviour*, 79(6), 1353–1362.
<https://doi.org/10.1016/j.anbehav.2010.03.009>
- Estes, D. (1998). Young Children's Awareness of Their Mental Activity: The Case of Mental Rotation. *Child Development*, 69(5), 1345–1360. <https://doi.org/10.1111/j.1467-8624.1998.tb06216.x>
- Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, 8(3), 223–241.
<https://doi.org/10.1177/1745691612460685>
- Fay, N., De Kleine, N., Walker, B., & Caldwell, C. A. (2019). Increasing population size can inhibit cumulative cultural evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 116(14), 6726–6731.
<https://doi.org/10.1073/pnas.1811413116>
- Fisher, J., & Hinde, R.A. (1949). The opening of milk bottles by birds. *British Birds*, 42, 347–357.
- Flavell, J. H. (1979). Metacognition and Cognitive Monitoring: A New Area of Cognitive-Development Inquiry. *American Psychologist*, 34(10), 906–911.
<https://doi.org/10.1037/0003-066X.34.10.906>
- Flynn, E. (2008). Investigating children as cultural magnets: do young children transmit redundant information along diffusion chains? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363(1509), 3541–3551.

- <https://doi.org/10.1098/rstb.2008.0136>
- Flynn, E., & Whiten, A. (2008). Cultural transmission of tool use in young children: A diffusion chain study. *Social Development, 17*(3), 699–718.
<https://doi.org/10.1111/j.1467-9507.2007.00453.x>
- Frick, A., Daum, M. M., Walser, S., & Mast, F. W. (2009). Motor Processes in Children's Mental Rotation. *Journal of Cognition and Development, 10*(1–2), 18–40.
<https://doi.org/10.1080/15248370902966719>
- Frick, A., Hansen, M. A., & Newcombe, N. S. (2013). Development of mental rotation in 3- to 5-year-old children. *Cognitive Development, 28*(4), 386–399.
<https://doi.org/10.1016/j.cogdev.2013.06.002>
- Galef, B. G. (2013). Imitation and local enhancement: detrimental effects of consensus definitions on analyses of social learning in animals. *Behavioural Processes, 100*, 123–130. <https://doi.org/10.1016/j.beproc.2013.07.026>
- Garland, E. C., Goldizen, A. W., Rekdahl, M. L., Constantine, R., Garrigue, C., Hauser, N. D., Poole, M. M., Robbins, J., & Noad, M. J. (2011). Dynamic horizontal cultural transmission of humpback whale song at the ocean basin scale. *Current Biology, 21*(8), 687–691. <https://doi.org/10.1016/j.cub.2011.03.019>
- Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31–60.
<https://doi.org/10.1037/0033-2909.134.1.31>
- Hampton, R. R., Zivin, A., & Murray, E. A. (2004). Rhesus monkeys (*Macaca mulatta*) discriminate between knowing and not knowing and collect information as needed before acting. *Animal Cognition, 7*(4), 239–246. <https://doi.org/10.1007/s10071-004-0215-1>
- Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour, 59*(4), 771–785.
<https://doi.org/10.1006/anbe.1999.1377>
- Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour, 61*(1), 139–151. <https://doi.org/10.1006/anbe.2000.1518>
- Haun, D. B. M., Rekers, Y., & Tomasello, M. (2012). Majority-biased transmission in chimpanzees and human children, but not orangutans. *Current Biology, 22*(8), 727–731. <https://doi.org/10.1016/j.cub.2012.03.006>
- Haun, D. B. M., & Tomasello, M. (2011). Conformity to Peer Pressure in Preschool Children. *Child Development, 82*(6), 1759–1767. <https://doi.org/10.1111/j.1467-8624.2011.01666.x>
- Henrich, J. (2004). Demography and Cultural Evolution: How Adaptive Cultural Processes

- Can Produce Maladaptive Losses—The Tasmanian Case. *American Antiquity*, 69(2), 197-214. doi:10.2307/4128416
- Henrich, J., Boyd, R., Derex, M., Kline, M. A., Mesoudi, A., Muthukrishna, M., Powell, A. T., Shennan, S. J., & Thomas, M. G. (2016). Understanding cumulative cultural evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 113(44), E6724–E6725. <https://doi.org/10.1073/pnas.1610005113>
- Henrich, J., Boyd, R., & Richerson, P. J. (2008). Five misunderstandings about cultural evolution. *Human Nature*, 19(2), 119–137. <https://doi.org/10.1007/s12110-008-9037-1>
- Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, 317(5843), 1360–1366. <https://doi.org/10.1126/science.1146282>
- Heyes, C. (1994). Social learning in animals: Categories and mechanisms. *Biology Review*, 69, 207–231.
- Heyes, C. (2012). What's social about social learning? *Journal of Comparative Psychology*, 126(2), 193–202. <https://doi.org/10.1037/a0025180>
- Heyes, C. (2015). Animal mindreading: what's the problem? *Psychonomic Bulletin and Review*, 22(2), 313–327. <https://doi.org/10.3758/s13423-014-0704-4>
- Heyes, C. (2016). Who Knows? Metacognitive Social Learning Strategies. *Trends in Cognitive Sciences*, 20(3), 204–213. <https://doi.org/10.1016/j.tics.2015.12.007>
- Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Harvard University Press. <https://doi.org/10.4159/9780674985155>
- Hobaiter, C., Poisot, T., Zuberbühler, K., Hoppitt, W., & Gruber, T. (2014). Social Network Analysis Shows Direct Evidence for Social Transmission of Tool Use in Wild Chimpanzees. *PLoS Biology*, 12(9). <https://doi.org/10.1371/journal.pbio.1001960>
- Hunt, G. R., & Gray, R. D. (2003). Diversification and cumulative evolution in New Caledonian crow tool manufacture. *Proceedings of the Royal Society B: Biological Sciences*, 270(1517), 867–874. <https://doi.org/10.1098/rspb.2002.2302>
- Iachini, T., Ruggiero, G., Bartolo, A., Rapuano, M., & Ruotolo, F. (2019). The Effect of Body-Related Stimuli on Mental Rotation in Children, Young and Elderly Adults. *Scientific Reports*, 9(1), 3–12. <https://doi.org/10.1038/s41598-018-37729-7>
- Jacobs, R. C., & Campbell, D. T. (1961). The perpetuation of an arbitrary tradition through several generations of a laboratory microculture. *The Journal of Abnormal and Social Psychology*, 62(3), 649–658. <https://doi.org/10.1037/h0044182>
- Jesmer, B. R., Merkle, J. A., Goheen, J. R., Aikens, E. O., Beck, J. L., Courtemanch, A. B., Hurley, M. A., McWhirter, D. E., Miyasaki, H. M., Monteith, K. L., & Kauffman, M. J. (2018). Is ungulate migration culturally transmitted? Evidence of social learning from

- translocated animals. *Science*, 361(6406), 1023–1025.
<https://doi.org/10.1126/science.aat0985>
- Jurowski, K., Jurowska, A., Krzeczowska, M., Jurowski, K., Jurowska, A., & Krzeczowska, M. (2015). Comprehensive review of mnemonic devices and their applications : state of the art. *International E-Journal of Science, Medicine and Education*, 9(3), 4–9.
- Kail, R., Pellegrino, J., & Carter, P. (1980). Developmental changes in mental rotation. *Journal of Experimental Child Psychology*, 29(1), 102–116.
[https://doi.org/10.1016/0022-0965\(80\)90094-6](https://doi.org/10.1016/0022-0965(80)90094-6)
- Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition*, 109(2), 224–234.
<https://doi.org/https://doi.org/10.1016/j.cognition.2008.08.010>
- Kapitány, R., Davis, J. T., Legare, C., & Nielsen, M. (2018). An experimental examination of object-directed ritualized action in children across two cultures. *PLoS ONE*, 13(11), 1–15. <https://doi.org/10.1371/journal.pone.0206884>
- Kempe, M., & Mesoudi, A. (2014). An experimental demonstration of the effect of group size on cultural accumulation. *Evolution and Human Behavior*, 35(4), 285–290.
<https://doi.org/10.1016/j.evolhumbehav.2014.02.009>
- Kendal, R. L., Boogert, N. J., Rendell, L., Laland, K. N., Webster, M., & Jones, P. L. (2018). Social Learning Strategies: Bridge-Building between Fields. *Trends in Cognitive Sciences*, 22(7), 651–665. <https://doi.org/10.1016/j.tics.2018.04.003>
- Kline, M. A., & Boyd, R. (2010). Population size predicts technological complexity in Oceania. *Proceedings of the Royal Society B: Biological Sciences*, 277(1693), 2559–2564. <https://doi.org/10.1098/rspb.2010.0452>
- Kloo, D., Rohwer, M., & Perner, J. (2017). Direct and indirect admission of ignorance by children. *Journal of Experimental Child Psychology*, 159, 279–295.
<https://doi.org/10.1016/j.jecp.2017.02.014>
- Kobayashi, Y., & Aoki, K. (2012). Innovativeness, population size and cumulative cultural evolution. *Theoretical Population Biology*, 82(1), 38–47.
<https://doi.org/10.1016/j.tpb.2012.04.001>
- Köhler, C., Hoffmann, K. P., Dehnhardt, G., & Mauck, B. (2005). Mental rotation and rotational invariance in the Rhesus monkey (*Macaca mulatta*). *Brain, Behavior and Evolution*, 66(3), 158–166. <https://doi.org/10.1159/000087156>
- Krachun, C., Carpenter, M., Call, J., & Tomasello, M. (2009). A competitive nonverbal false belief task for children and apes. *Developmental Science*, 12(4), 521–535.
<https://doi.org/10.1111/j.1467-7687.2008.00793.x>
- Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate

- that other individuals will act according to false beliefs. *Science (New York, N. Y.)*, 354(6308), 110–114. <https://doi.org/10.1126/science.aaf8110>
- Lake, M.W., & Venti, J. (2009) Quantitative analysis of macroevolutionary patterning in technological evolution: Bicycle design from 1800 to 2000. In S.J. Shennan (Ed.) *Pattern and Process in Cultural Evolution*. (pp. 147-174). University of California Press.
- Laland, K. N. (2004). Social learning strategies. *Learning and Behavior*, 32(1), 4–14. <https://doi.org/10.3758/bf03196002>
- Laland, K. N., & Galef, B. G. (Eds.). (2009). *The question of animal culture*. Harvard University Press.
- Laland, K. N., & Hoppitt, W. (2003). Do Animals Have Culture? *Evolutionary Anthropology*, 12(3), 150–159. <https://doi.org/10.1002/evan.10111>
- Leadbeater, E., & Chittka, L. (2007). The dynamics of social learning in an insect model, the bumblebee (*Bombus terrestris*). *Behavioral Ecology and Sociobiology*, 61(11), 1789–1796. <https://doi.org/10.1007/s00265-007-0412-4>
- Legare, C. H., & Nielsen, M. (2015). Imitation and Innovation: The Dual Engines of Cultural Learning. *Trends in Cognitive Sciences*, 19(11), 688–699. <https://doi.org/10.1016/j.tics.2015.08.005>
- Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology*, 35(4), 940–949. <https://doi.org/10.1037/0012-1649.35.4.940>
- Lewis, H. M., & Laland, K. N. (2012). Transmission fidelity is the key to the build-up of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1599), 2171–2180. <https://doi.org/10.1098/rstb.2012.0119>
- Luncz, L. V., & Boesch, C. (2014). Tradition over trend: Neighboring chimpanzee communities maintain differences in cultural behavior despite frequent immigration of adult females. *American Journal of Primatology*, 76(7), 649–657. <https://doi.org/10.1002/ajp.22259>
- McGuigan, N., Burdett, E., Burgess, V., Dean, L., Lucas, A., Vale, G., & Whiten, A. (2017). Innovation and social transmission in experimental micro-societies: Exploring the scope of cumulative culture in young children. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1735). <https://doi.org/10.1098/rstb.2016.0425>
- Menzel E.W., J., Davenport, R. K., & Rogers, C. M. (1972). Protocultural Aspects of Chimpanzees' Responsiveness to Novel Objects. *Folia Primatologica*, 17(3), 161–170. <https://doi.org/10.1159/000155425>
- Mesoudi, A., & Whiten, A. (2008). The multiple roles of cultural transmission experiments in

- understanding human cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1509), 3489–3501.
<https://doi.org/10.1098/rstb.2008.0129>
- Mesoudi, Alex, & Thornton, A. (2018). What is cumulative cultural evolution? *Proceedings of the Royal Society B: Biological Sciences*, 285(1880).
<https://doi.org/10.1098/rspb.2018.0712>
- Moll, H., & Tomasello, M. (2006). Level I perspective-taking at 24 months of age. *British Journal of Developmental Psychology*, 24(3), 603–613.
<https://doi.org/10.1348/026151005X55370>
- Morgan, T. J. H., Rendell, L. E., Ehn, M., Hoppitt, W., & Laland, K. N. (2012). The evolutionary basis of human social learning. *Proceedings of the Royal Society B: Biological Sciences*, 279(1729), 653–662. <https://doi.org/10.1098/rspb.2011.1172>
- Morgan, T. J. H., Uomini, N. T., Rendell, L. E., Chouinard-Thuly, L., Street, S. E., Lewis, H. M., Cross, C. P., Evans, C., Kearney, R., De La Torre, I., Whiten, A., & Laland, K. N. (2015). Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nature Communications*, 6, 4–11. <https://doi.org/10.1038/ncomms7029>
- Muthukrishna, M., & Henrich, J. (2016). Innovation in the collective brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1690).
<https://doi.org/10.1098/rstb.2015.0192>
- Muthukrishna, M., Shulman, B. W., Vasilescu, V., & Henrich, J. (2014). Sociality influences cultural complexity. *Proceedings. Biological Sciences*, 281(1774), 20132511.
<https://doi.org/10.1098/rspb.2013.2511>
- Nakamichi, K. (2019). Young children’s counterfactual thinking: Triggered by the negative emotions of others. *Journal of Experimental Child Psychology*, 187, 104659.
<https://doi.org/10.1016/j.jecp.2019.06.012>
- Neldner, K., Mushin, I., & Nielsen, M. (2017). Young children’s tool innovation across culture: Affordance visibility matters. *Cognition*, 168, 335–343.
<https://doi.org/10.1016/j.cognition.2017.07.015>
- O’Brien, M. J., & Lyman, R. L. (2000). Darwinian Evolutionism Is Applicable to Historical Archaeology. *International Journal of Historical Archaeology*, 4(1), 71–112.
<https://doi.org/10.1023/A:1009556427520>
- O’Neill, D. K. (1996). Two-Year-Old Children’s Sensitivity to a Parent’s Knowledge State When Making Requests. *Child Development*, 67(2), 659–677.
<https://doi.org/https://doi.org/10.1111/j.1467-8624.1996.tb01758.x>
- O’Neill, D. K., Astington, J. W., & Flavell, J. H. (1992). Young Children’s Understanding of the Role That Sensory Experiences Play in Knowledge Acquisition. *Child Development*,

- 63(2), 474–490. <https://doi.org/10.1111/j.1467-8624.1992.tb01641.x>
- Osiurak, F., De Oliveira, E., Navarro, J., Lesourd, M., Claidière, N., & Reynaud, E. (2016). Physical intelligence does matter to cumulative technological culture. *Journal of Experimental Psychology: General*, *145*(8), 941–948. <https://doi.org/10.1037/xge0000189>
- Osiurak, F., De Oliveira, E., Navarro, J., & Reynaud, E. (2020). The castaway island: Distinct roles of theory of mind and technical reasoning in cumulative technological culture. *Journal of Experimental Psychology: General*, *149*(1), 58–66. <https://doi.org/10.1037/xge0000614>
- Osiurak, F., & Reynaud, E. (2020). The Elephant in the Room: What Matters Cognitively in Cumulative Technological Culture. *Behavioral and Brain Sciences*, 1–57. <https://doi.org/10.1017/s0140525x19003236>
- O’Sullivan, E.P. (2015). *A comparative approach to social learning from the bottom up* [Unpublished doctoral dissertation]. University of Stirling.
- Palmquist, C. M., Burns, H. E., & Jaswal, V. K. (2012). Pointing disrupts preschoolers’ ability to discriminate between knowledgeable and ignorant informants. *Cognitive Development*, *27*(1), 54–63. <https://doi.org/10.1016/j.cogdev.2011.07.002>
- Pearce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Pike, T. W., & Laland, K. N. (2010). Conformist learning in nine-spined sticklebacks’ foraging decisions. *Biology Letters*, *6*(4), 466–468. <https://doi.org/10.1098/rsbl.2009.1014>
- Pillow, B. H. (1989). Early understanding of perception as a source of knowledge. *Journal of Experimental Child Psychology*, *47*(1), 116–129. [https://doi.org/10.1016/0022-0965\(89\)90066-0](https://doi.org/10.1016/0022-0965(89)90066-0)
- Povinelli, D. J., & DeBlois, S. (1992). Young children's (Homo sapiens) understanding of knowledge formation in themselves and others. *Journal of Comparative Psychology*, *106*(3), 228–238. <https://doi.org/10.1037/0735-7036.106.3.228>
- Powell, A., Shennan, S., & Thomas, M. G. (2009). Late pleistocene demography and the appearance of modern human behavior. *Science*, *324*(5932), 1298–1301. <https://doi.org/10.1126/science.1170165>
- Pradhan, G. R., Tennie, C., & van Schaik, C. P. (2012). Social organization and the evolution of cumulative technology in apes and hominins. *Journal of Human Evolution*, *63*(1), 180–190. <https://doi.org/10.1016/j.jhevol.2012.04.008>
- Pratt, C., & Bryant, P. (1990). Young Children Understand That Looking Leads to Knowing (So Long as They Are Looking into a Single Barrel). *Child Development*, *61*(4), 973–

982. <https://doi.org/10.1111/j.1467-8624.1990.tb02835.x>
- Premack, D., & Woodruff, G. (1978). Premack and Woodruff : Chimpanzee theory of mind. *Behavioral and Brain Sciences*, 4(1978), 515–526.
- R Core Team (2018). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. URL <https://www.R-project.org/>.
- Rafetseder, E., Cristi-Vargas, R., & Perner, J. (2010). Counterfactual reasoning: Developing a sense of “nearest possible world.” *Child Development*, 81(1), 376–389. <https://doi.org/10.1111/j.1467-8624.2009.01401.x>
- Rafetseder, E., & Perner, J. (2014). Counterfactual reasoning: Sharpening conceptual distinctions in developmental studies. *Child Development Perspectives*, 8(1), 54–58. <https://doi.org/10.1111/cdep.12061>
- Rafetseder, E., Schwitalla, M., & Perner, J. (2013). Counterfactual reasoning: From childhood to adulthood. *Journal of Experimental Child Psychology*, 114(3), 389–404. <https://doi.org/10.1016/j.jecp.2012.10.010>
- Reader, S. M., & Biro, D. (2010). Experimental identification of social learning in wild animals. *Learning & Behavior*, 38(3), 265–283. <https://doi.org/10.3758/LB.38.3.265>
- Reindl, E., Gwilliams, A. L., Dean, L. G., Kendal, R. L., & Tennie, C. (2020). Skills and motivations underlying children’s cumulative cultural learning: case not closed. *Palgrave Communications*, 6(1), 1–9. <https://doi.org/10.1057/s41599-020-0483-7>
- Reindl, E., & Tennie, C. (2018). Young children fail to generate an additive ratchet effect in an open-ended construction task. *PLoS ONE*, 13(6), 1–22. <https://doi.org/10.1371/journal.pone.0197828>
- Rendell, L., Fogarty, L., Hoppitt, W. J. E., Morgan, T. J. H., Webster, M. M., & Laland, K. N. (2010). Cognitive culture: Theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences*, 15(2), 68–76. <https://doi.org/10.1016/j.tics.2010.12.002>
- Renner, E., Kean, D., Atkinson, M., & Caldwell, C.A. (submitted). Capuchin monkeys’ (*Sapajus apella*) and human children’s (*Homo sapiens*) use of information from social, virtual, and individual sources in a stimulus choice task.
- Richerson, P. (2013). Human evolution: Group size determines cultural complexity. *Nature*, 503(7476), 351–352. <https://doi.org/10.1038/nature12708>
- Robinson, E. J., Rowley, M. G., Beck, S. R., Carroll, D. J., & Apperly, I. A. (2006). Children’s sensitivity to their own relative ignorance: handling of possibilities under epistemic and physical uncertainty. *Child Development*, 77(6), 1642–1655. <https://doi.org/10.1111/j.1467-8624.2006.00964.x>
- Roebbers, C. M. (2017). Executive function and metacognition: Towards a unifying framework

- of cognitive self-regulation. *Developmental Review*, 45, 31–51.
<https://doi.org/10.1016/j.dr.2017.04.001>
- Rohwer, M., Kloo, D., & Perner, J. (2012). Escape From Metaignorance: How Children Develop an Understanding of Their Own Lack of Knowledge. *Child Development*, 83(6), 1869–1883. <https://doi.org/10.1111/j.1467-8624.2012.01830.x>
- Sarin, S., & Dukas, R. (2009). Social learning about egg-laying substrates in fruitflies. *Proceedings of the Royal Society B: Biological Sciences*, 276(1677), 4323–4328.
<https://doi.org/10.1098/rspb.2009.1294>
- Sasaki, T., & Biro, D. (2017). Cumulative culture can emerge from collective intelligence in animal groups. *Nature Communications*, 8, 1–6. <https://doi.org/10.1038/ncomms15049>
- Schmidt, E., & Pyers, J. (2011). Children's Understanding of the Link Between Sensory Perception and Knowledge. In L. Carlson, C. Hoelscher, & T. Shipley (Eds), *The Proceedings of the 33rd Annual Meeting of the Cognitive Science Society* (pp. 3016-21). Austin: Cognitive Science Society
- Schofield, D. P., McGrew, W. C., Takahashi, A., & Hirata, S. (2017). Cumulative culture in nonhumans: overlooked findings from Japanese monkeys? *Primates*, 59(2), 113–122.
<https://doi.org/10.1007/s10329-017-0642-7>
- Shennan, S. (2001). Demography and cultural innovation: A model and its implications for the emergence of modern human culture. *Cambridge Archaeological Journal*, 11(1), 5–16. <https://doi.org/10.1017/s0959774301000014>
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171(3972), 701–703.
- Smith, J. D., Shields, W. E., Schull, J., & Washburn, D. A. (1997). The uncertain response in humans and animals. *Cognition*, 62(1), 75–97. [https://doi.org/10.1016/S0010-0277\(96\)00726-3](https://doi.org/10.1016/S0010-0277(96)00726-3)
- Sodian, B., & Wimmer, H. (1987). Children's understanding of inference as a source of knowledge. *Child Development*, 58(2), 424–433. <https://doi.org/10.2307/1130519>
- Sophian, C., & Somerville, S. C. (1988). Early developments in logical reasoning: Considering alternative possibilities. *Cognitive Development*, 3(2), 183–222.
[https://doi.org/10.1016/0885-2014\(88\)90018-4](https://doi.org/10.1016/0885-2014(88)90018-4)
- Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, 18(7), 587–592.
<https://doi.org/10.1111/j.1467-9280.2007.01944.x>
- Stich, K. P., Dehnhardt, G., & Mauck, B. (2003). Mental rotation of perspective stimuli in a California sea lion (*Zalophus californianus*). *Brain, Behavior and Evolution*, 61(2), 102–112. <https://doi.org/10.1159/000069355>

- Subiaul, F., Krajkowski, E., Price, E. E., & Etz, A. (2015). Imitation by combination: preschool age children evidence summative imitation in a novel problem-solving task. *Frontiers in Psychology*, 6(September), 1–13. <https://doi.org/10.3389/fpsyg.2015.01410>
- Taylor, M. (1988). *The development of children's understanding of the seeing-knowing distinction*. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (p. 207–225). Cambridge University Press.
- Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: On the evolution of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1528), 2405–2415. <https://doi.org/10.1098/rstb.2009.0052>
- Tennie, C., Walter, V., Gampe, A., Carpenter, M., & Tomasello, M. (2014). Limitations to the cultural ratchet effect in young children. *Journal of Experimental Child Psychology*, 126, 152–160. <https://doi.org/10.1016/j.jecp.2014.04.006>
- Terkel, J. (1996). *Cultural transmission of feeding behavior in the black rat (Rattus rattus)*. In C. M. Heyes & B. G. Galef, Jr. (Eds.), *Social learning in animals: The roots of culture* (p. 17–47). Academic Press. <https://doi.org/10.1016/B978-012273965-1/50003-0>
- Thornton, A., & McAuliffe, K. (2006). Teaching in Wild Meerkats. *Science*, 313(5784), 227–229. <https://doi.org/10.1126/science.1128727>
- Tomasello, M. (1990). Cultural transmission in the tool use and communicatory signaling of chimpanzees? In S. T. Parker & K. R. Gibson (Eds.), *“Language” and intelligence in monkeys and apes* (pp. 274–311). Cambridge University Press. <https://doi.org/10.1017/CBO9780511665486.012>
- Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.
- Tomasello, M., Call, J., Nagell, K., Olguin, R., & Carpenter, M. (1994). The learning and use of gestural signals by young chimpanzees: A trans-generational study. *Primates*, 35(2), 137–154. <https://doi.org/10.1007/BF02382050>
- Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, 16(3), 495–552. <https://doi.org/10.1017/s0140525x0003123x>
- Vaesen, K., Collard, M., Cosgrove, R., & Roebroeks, W. (2016). Population size does not explain past changes in cultural complexity. *Proceedings of the National Academy of Sciences of the United States of America*, 113(16), E4241–E4247. <https://doi.org/10.1073/pnas.1520288113>
- Van Leeuwen, E. J. C., Cohen, E., Collier-Baker, E., Rapold, C. J., Schäfer, M., Schütte, S., & Haun, D. B. M. (2018). The development of human social learning across seven societies. *Nature Communications*, 9(1), 1–7. <https://doi.org/10.1038/s41467-018-04468-2>
- Vauclair, J., Fagot, J., & Hopkins, W. D. (1993). Rotation of Mental Images in Baboons

- When the Visual Input Is Directed to the Left Cerebral Hemisphere. *Psychological Science*, 4(2), 99–103. <https://doi.org/10.1111/j.1467-9280.1993.tb00468.x>
- Wasielewski, H. (2014). Imitation Is Necessary for Cumulative Cultural Evolution in an Unfamiliar, Opaque Task. *Human Nature*, 25(1), 161–179. <https://doi.org/10.1007/s12110-014-9192-5>
- Wellman, H. M., Cross, D., & Watson, J. (2001). *Meta-Analysis of Theory-of-Mind Development : The Truth about False Belief*. 72(3), 655–684.
- Whiten, A., McGuigan, N., Marshall-Pescini, S., & Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1528), 2417–2428. <https://doi.org/10.1098/rstb.2009.0069>
- Whiten, A., & van Schaik, C. P. (2007). The evolution of animal “cultures” and social intelligence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480), 603–620. <https://doi.org/10.1098/rstb.2006.1998>
- Whiten, A, Custance, D. M., Gomez, J.-C., Teixidor, P., & Bard, K. A. (1996). Imitative learning of artificial fruit-processing in children (*Homo sapiens*) and chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.*, 110(1), 3–14.
- Whiten, A. (2013). Humans are not alone in computing how others see the world. *Animal Behaviour*, 86(2), 213–221. <https://doi.org/10.1016/j.anbehav.2013.04.021>
- Whiten, A., Caldwell, C. A., & Mesoudi, A. (2016). Cultural diffusion in humans and other animals. *Current Opinion in Psychology*, 8, 15–21. <https://doi.org/10.1016/j.copsyc.2015.09.002>
- Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C. E. G., Wrangham, R. W., & Boesch, C. (1999). Cultures in chimpanzees. *Nature*, 399(6737), 682–685. <https://doi.org/10.1038/21415>
- Whiten, A., McGuigan, N., Marshall-pescini, S., & Hopper, L. M. (2009). *Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee*. 2417–2428. <https://doi.org/10.1098/rstb.2009.0069>
- Wilks, C. E. H., & Blakey, K. H. (2018). In the jungle of cultural complexity. *Evolutionary Anthropology*, 27(5), 180–183. <https://doi.org/10.1002/evan.21724>
- Wilks, C.E.H., Rafetseder, E., Renner, E., Atkinson, M., & Caldwell, C.A. (in press). Cognitive Prerequisites for Cumulative Culture are Context-Dependent: Children’s Potential for Ratcheting Depends on Cue Longevity. *J Exp Child Psychol*.
- Wilks, M., Collier-Baker, E., & Nielsen, M. (2015). Preschool children favor copying a successful individual over an unsuccessful group. *Developmental Science*, 18(6), 1014–1024. <https://doi.org/10.1111/desc.12274>

- Wimmer, H., Hogrefe, G.J., & Perner, J. (1988). Children's understanding of informational access as source of knowledge. *Child Development*, 59(2), 386–396. <https://doi.org/10.2307/1130318>
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128. [https://doi.org/10.1016/0010-0277\(83\)90004-5](https://doi.org/10.1016/0010-0277(83)90004-5)
- Winters, J. (2020, September 26). Is the cultural evolution of technology cumulative or combinatorial?. <https://doi.org/10.31235/osf.io/aypnx>
- Wood, L. A., Kendal, R. L., & Flynn, E. G. (2013). Whom do children copy? Model-based biases in social learning. *Developmental Review*, 33(4), 341–356. <https://doi.org/10.1016/j.dr.2013.08.002>
- Wynn, T., & Coolidge, F. L. (2011). The Implications of the Working Memory Model for the Evolution of Modern Cognition. *International Journal of Evolutionary Biology*, 2011, 1–12. <https://doi.org/10.4061/2011/741357>
- Zwirner, E., & Thornton, A. (2015). Cognitive requirements of cumulative culture: Teaching is useful but not essential. *Scientific Reports*, 5, 1–8. <https://doi.org/10.1038/srep16781>

Appendix

Supplementary Information – Chapter 2

Materials and Methods

Participants – Glasgow Science Centre. Glasgow Science Centre is a science museum with an entry fee in Glasgow, UK. Three children were excluded due to: missing date of birth from the consent form (age 3, female), failure to fully comply with task instructions (age 4, male), and experimenter error (age 6, female). The final sample from Glasgow Science Centre consisted of 80 children aged 3-6 (M= 59 months, range= 38 – 83, SD= 13, 50 female). 40 were assigned to each condition.

Participants – School. The Ofsted (Office for Standards in Education) inspection report for the school we tested at in Bradford stated that the proportion of pupils for which pupil premium (funding for children from low-income families) is received is just below average. Children were recruited from the following year-groups: nursery (aged 3-4), reception (aged 4-5), year one (aged 5-6), and Year 2 (aged 6, and one aged 7). When an age category spanned two school year-groups (all except age 3), we tried to ensure our sample was a 50:50 split between the two year-groups in order to control for the effects of schooling. Six children were excluded for the following reasons: inability to understand the task due to language difficulty (age 3, male), failure to fully comply with task instructions (n=3; aged 3, 4 and 5; male), and experimenter error (n=2, age 4, female). The final sample from the primary school consisted of 74 children aged 3-7 (M= 59 months, range= 38 – 84, SD= 13, 38 female). 36 and 38 children were assigned to the Enduring Cues and Transient Information conditions respectively.

Results

Optimal Response Count. To use the social information optimally, the puppet's chest selections should be repeated only if rewarded; if unrewarded an alternative chest should be selected by the child. Adopting this behaviour should maximise the amount of treasure found. Hence, for each child we calculated an optimal response count to represent how well they were using both rewarded and unrewarded information i.e. their proficiency using the social information overall. In our analysis of children's optimal response counts, we were interested in the effect of age (3-6), condition (Transient Information or Enduring Cues) and the interaction of age and condition on optimal response count.

Repeating. The optimal response count allowed overall examination of rewarded

and unrewarded social information use but did not allow the responses to rewarded and unrewarded selections to be differentiated. Consequently, we analysed the children's responses in terms of "repeating" after both rewarded and unrewarded puppet selections. Considering repeating also allowed us to analyse how the use of rewarded and unrewarded information may differ according to condition and age (Table 2.1; Figures 2.3A and 2.3B). We were interested in the effect of information type (rewarded or unrewarded puppet selection), age (3-6), condition (Transient Information or Enduring Cues), and the interaction of these three variables on tendency to repeat the selection performed by the puppet.

Potential for Ratcheting. We used one-sample (one-tailed) t-tests when assessing whether mean success scores (Table S.1 and Figure 2.4) were significantly higher than chance level (success score 3/9 or 1/3), and 2/3 demonstration types. We also used one-sample (one-tailed) t-tests when testing whether scores were higher, on average, following higher success (2/3 and 3/3) demonstrations compared with lower (1/3) demonstrations e.g. to test whether the mean score for the 2/3 and 3/3 demonstration types (Table S.1) was significantly above 2 (theoretically optimal performance following a 1/3 demonstration). We used one-tailed tests because we were only interested in testing whether success scores exceeded the benchmarks outlined in our PFR level criteria.

In the Enduring Cues condition, and according to the strict criteria, the mean success scores (out of nine for the entire task) were significantly above chance-level performance (a score of three) at each age group (3-6), all p values $<.001$ (Table S.1 and Figure 2.4). This meant that all age groups were above the criterion for Level 1. Moreover, the mean success scores for the 2/3 and 3/3 demonstration types combined were significantly above 2 at each age group, all p values $<.001$ (Table S.1), therefore all age groups reached the minimum requirement for Level 2. Importantly, the mean success scores also showed outperformance of the puppet's score for both the 1/3 and 2/3 rewarded arrays (scores of significantly above one and two respectively) but only for ages 4 and 6 (Table S.1 and Figure 2.4). We therefore classified children aged 4 and 6 only as showing Level 3 ratcheting: we would expect ratcheting of the information presented to occur in a linear transmission chain. Children aged 3 and 5 were categorised as Level 2 which showed that, although they were not able to outperform the demonstrations, they were able to score more highly following higher success demonstrations (Figure 2.5).

In contrast, in the Transient Information condition (again according to the strict criteria) we did not observe Level 3 ratcheting, consistent with improved performance in a linear transmission chain, at any age group (Figure 2.5) because there was no significant outperformance of the 2/3 demonstration, all p values $>.05$ (Table S.1 and Figure 2.4).

However, the mean success scores, out of nine for the entire task, were again significantly above chance-level performance (a score of three) at each age group (3-6) and all were therefore above the Level 1 criterion, all p values $<.01$ (Table S.1). Yet, the mean success scores for the 2/3 and 3/3 demonstration types combined were not significantly above 2 at any age group, all p values $>.05$ (Table S.1 and Figure 2.4), and children aged 3-6 were therefore all categorised as Level 1 (Figure 2.5).

Memory Question. We recorded children's responses to the memory question: "Where did Polly the Parrot look for the [red] treasure?" These were coded as "1" for a correct response (child correctly stated/pointed at the location selected by the puppet) and "0" for an incorrect response (child indicated the incorrect location). This memory question was asked in the Transient Information condition only and all children responded.

Responses to the memory question allowed us to investigate how the different cognitive constraints presented by our Enduring Cues and Transient Information conditions affected use of the social information. It is likely that the demands the Transient Information condition placed on memory led to the age effects observed in this condition in our optimal response count and repeating analyses (see page 52 "Effects of Information Type and Age in the Enduring Cues and Transient Information Conditions"), and thus the lower level of PFR classification.

In Table S.2 we display the mean (of participant means) proportion of locations remembered which increased with age for both the unrewarded and rewarded selections but to a greater extent for the rewarded selections. Interestingly, on page 54 (Effects of Age and Condition by Information Type) we showed increased repeating with age for rewarded selections, but not decreased repeating for unrewarded selections which may have been due to this increased memory for the locations selected in rewarded trials, compared to unrewarded trials, with age.

Table S.1

Mean, Standard Deviation and One-Sample T-Tests (One-Tailed) for Success Score/9 and Success Score/3 for Ages 3-6 (Whole Years) in the Transient Information and Enduring Cues Conditions

Note. One-sample t-tests were used to determine if overall success scores/9 were significantly higher than 3 and success scores/3 were significantly higher than the following for the different array types: 1 (1 rewarded array) and 2 (2 rewarded and the mean of 2 and 3 rewarded arrays).

Age	Condition	Success score/9				Success score/3													
		Mean	SD	t	p	1 rewarded				2 rewarded				3 rewarded		2 and 3 rewarded			
						Mean	SD	t	p	Mean	SD	t	p	Mean	SD	Mean	SD	t	p
3	Enduring	6.41	1.33	10.611	<.001	1.71	0.92	3.165	.003	2.06	0.90	0.270	.395	2.65	0.61	2.35	0.52	2.781	.007
4	Enduring	6.75	1.29	12.973	<.001	1.75	0.72	4.682	<.001	2.30	0.66	2.042	.028	2.70	0.73	2.50	0.46	4.873	<.001
5	Enduring	6.65	1.35	12.105	<.001	1.75	0.72	4.682	<.001	2.15	0.93	0.719	.241	2.75	0.55	2.45	0.54	3.758	<.001
6	Enduring	7.42	0.90	21.374	<.001	2.11	0.74	6.533	<.001	2.42	0.61	3.024	.004	2.89	0.32	2.66	0.34	8.547	<.001
3	Transient	3.90	1.37	2.932	.004	1.35	0.81	1.926	.035	1.45	0.89	-2.773	.994	1.10	0.97	1.28	0.55	-5.900	1.000
4	Transient	4.37	1.83	3.256	.002	1.37	0.76	2.111	.025	1.63	0.90	-1.794	.955	1.37	1.07	1.50	0.75	-2.924	.996
5	Transient	5.47	1.31	8.251	<.001	1.68	0.82	3.637	<.001	1.47	1.02	-2.249	.981	2.32	0.82	1.89	0.59	-0.776	.776
6	Transient	5.70	1.26	9.578	<.001	1.40	0.60	2.990	.004	1.70	0.86	-1.552	.931	2.60	0.88	2.15	0.49	1.371	.093

Table S.2

Mean Proportion of Locations Selected by the Puppet Which Were Remembered and Standard Deviation for all Ages Combined and Ages 3-6 (Whole Years) for Rewarded, Unrewarded and Rewarded/Unrewarded Information Combined

Age (whole years)	Information type	Mean proportion locations remembered	SD
3	All	.39	0.49
4	All	.56	0.50
5	All	.60	0.49
6	All	.73	0.45
All	Unrewarded	.52	0.50
All	Rewarded	.60	0.49
3	Unrewarded	.37	0.49
4	Unrewarded	.54	0.50
5	Unrewarded	.56	0.50
6	Unrewarded	.60	0.49
3	Rewarded	.41	0.49
4	Rewarded	.56	0.50
5	Rewarded	.62	0.49
6	Rewarded	.79	0.41

Table S.3 Final Structure of all GLM and GLMM Used in the Analysis of Optimal Response Count and Repetition of Responses

Final Model Structure (Maximal Random Variables)	Data Used	Fixed Effects	<i>b</i>	<i>SE</i>	<i>Z</i>	<i>p</i>	
Optimal Response Count ~ age_days.ct * condition	All	Intercept	1.944	0.031	63.231	< .001	***
		age_days.ct	0.204	0.077	2.658	.008	**
		condition	0.165	0.031	5.382	< .001	***
		age_days.ct:condition	0.142	0.077	-1.856	.063	
Optimal Response Count ~ age_days.ct	Enduring	Intercept	2.110	0.040	52.816	< .001	***
		age_days.ct	0.062	0.100	0.614	.539	
Optimal Response Count ~ age_days.ct	Transient	Intercept	1.774	0.047	37.842	< .001	***
		age_days.ct	0.347	0.116	2.981	.003	**
Repetition of Responses ~ age_days.ct * information_type * condition + (1+information_type participant) + (1 reward_position) + (1 trial_number)	All	Intercept	-0.474	0.126	-3.752	< .001	***
		age_days.ct	0.633	0.286	2.210	.027	*
		information_type	1.824	0.120	15.255	< .001	***
		condition	-0.193	0.116	-1.655	.098	
		age_days.ct:information_type	0.562	0.289	1.948	.051	
		age_days.ct:condition	0.131	0.286	0.457	.648	
		information_type:condition	1.003	0.117	8.579	< .001	***
Repetition of Responses ~ age_days.ct * information_type + (1+information_type participant) + (1 reward_position) + (1 trial_number)	Enduring	Intercept	-2.254	0.567	3.973	< .001	***
		age_days.ct	0.831	0.988	0.841	.400	
		information_type	4.824	0.524	9.214	< .001	***
		age_days.ct:information_type	0.286	0.972	0.295	.768	
Repetition of Responses ~ age_days.ct information_type + (1 participant) + (1 trial_number) + (1 remembered)	Transient	Intercept	-0.595	1.021	-0.583	.560	
		age_days.ct	-0.202	0.280	-0.723	.470	
		information_type	0.997	0.113	8.827	< .001	***
		age_days.ct:information_type	0.894	0.271	3.305	< .001	***
Repetition of Responses ~ age_days.ct * condition + (1 trial_number)	Rewarded	Intercept	1.338	0.099	13.498	< .001	***
		age_days.ct	1.184	0.238	4.979	< .001	***
		condition	0.804	0.093	8.646	< .001	***
		age_days.ct:condition	-0.187	0.238	-0.788	.431	
Repetition of Responses ~ age_days.ct * condition	Unrewarded	Intercept	-2.292	0.211	10.842	< .001	***
		age_days.ct	0.075	0.519	0.144	.885	
		condition	-1.194	0.211	5.646	< .001	***
		age_days.ct:condition	0.459	0.519	0.884	.377	

Figures**Figure S.1**

Photo Showing the Arrangement of Treasure Chests on the Laminated Treasure Map, Using Array 1 as an Example



Figure S.2

Photo of the Experimental Setup, Using Array 3 as an Example



Supplementary Information – Chapter 3

Results

Rewarded Repeats – Effects of Age by Condition. We ran four further GLMMs (one for each condition), with a view to understanding the differing effects of age on rewarded repeats in each of the four conditions (i.e. the interaction between age and condition). We split the data by condition and removed this from the original model. The current models were therefore identical to the first model (pp. 87-88) except that they had a fixed effect of age only. We found a main effect of age in the Temporal ($b = 0.099$, $SE = 0.039$, $Z = 2.565$, $p = 0.010$), Spatial ($b = 0.173$, $SE = 0.039$, $Z = 4.457$, $p < .001$) and Temporal-Spatial ($b = 0.232$, $SE = 0.041$, $Z = 5.633$, $p < .001$) conditions but not in the Increased Information condition ($b = 0.065$, $SE = 0.038$, $Z = 1.719$, $p = 0.086$). This showed that the number of rewarded repeats increased with increasing age in all conditions except the Increased Information (Figure 3.5A, Table 3.2).

Rewarded Repeats – Effects of Condition by Age. We ran six more GLMMs (one for age group), with a view to further understanding the differing effects of condition on rewarded repeats by age. We split the data by age and removed this from the original model therefore the current models were identical to the first model (pp. 87-88) except that they had a fixed effect of condition only. We found main effects of conditions Temporal-Spatial ($b = -0.364$, $SE = 0.093$, $Z = -3.915$, $p < .001$) and Spatial ($b = -0.230$, $SE = 0.089$, $Z = -2.575$, $p = 0.010$) in our model for children aged 5: fewer rewarded repeats in these conditions compared to the Increased Information condition (Figure 3.5A, Table 3.2). For children aged 6 there was a main effect of the Temporal-Spatial condition only ($b = -0.231$, $SE = 0.080$, $Z = -2.873$, $p = 0.004$): less rewarded repeats in this condition compared to the Increased Information. There were no main effects for the remaining conditions in children aged 5 and 6 and no main effects for any condition for children aged 7-10 (all $p > .080$): the number of rewarded repeats did not significantly differ from that in the Increased Information condition.

Table S.4

Mean, Standard Deviation and One-Sample T-Tests (One-Tailed) for the Mean Child Score/3 on the Model-Score-2 Trials – Ages 5-10 (Whole Years) in the Increased Information (I), Temporal (T), Spatial (S) and Temporal-Spatial (TS) Conditions

Age (whole years)	Condition	Model-score-2 trials							
		Child mean score/3		T-test – chance- level score (1.000)		T-test – mean score across both models (1.667)		T-test – highest single model score (2.000)	
		Mean	SD	t	p	t	p	t	p
5	I	2.369	0.803	9.022	< .001	4.626	< .001	2.432	.011
6	I	2.591	0.330	26.843	< .001	15.592	< .001	9.973	< .001
7	I	2.641	0.229	36.467	< .001	21.645	< .001	14.245	< .001
8	I	2.690	0.201	44.414	< .001	26.890	< .001	18.141	< .001
9	I	2.667	0.258	35.940	< .001	21.557	< .001	14.376	< .001
10	I	2.667	0.265	28.137	< .001	16.876	< .001	11.255	< .001
5	T	2.143	0.699	8.647	< .001	3.600	< .001	1.081	.145
6	T	2.452	0.568	14.230	< .001	7.692	< .001	4.427	< .001
7	T	2.641	0.187	44.809	< .001	26.596	< .001	17.504	< .001
8	T	2.631	0.344	25.121	< .001	14.848	< .001	9.719	< .001
9	T	2.656	0.279	33.085	< .001	19.758	< .001	13.105	< .001
10	T	2.583	0.457	15.486	< .001	8.963	< .001	5.706	< .001

5	S	1.786	0.704	5.906	< .001	0.892	0.190	-1.611	.941
6	S	2.419	0.447	17.681	< .001	9.372	< .001	5.224	< .001
7	S	2.500	0.464	16.474	< .001	9.149	< .001	5.491	< .001
8	S	2.667	0.157	56.125	< .001	33.664	< .001	22.450	< .001
9	S	2.645	0.271	33.765	< .001	20.075	< .001	13.241	< .001
10	S	2.717	0.196	39.226	< .001	23.985	< .001	16.376	< .001
5	TS	1.655	0.548	6.323	< .001	-0.118	0.547	-3.334	.999
6	TS	2.022	0.530	10.729	< .001	3.724	< .001	0.226	.411
7	TS	2.205	0.422	14.546	< .001	6.495	< .001	2.476	.010
8	TS	2.452	0.387	19.849	< .001	10.734	< .001	6.183	< .001
9	TS	2.505	0.309	27.145	< .001	15.118	< .001	9.113	< .001
10	TS	2.617	0.196	36.941	< .001	21.700	< .001	14.091	< .001

Note. One-sample t-tests were used to determine if the mean child score (out of 3) was significantly greater than the following: a chance-level score (1), the mean score across the two models (1.667) and/or the highest single model score (2).

Table S.5

Mean, Standard Deviation and One-Sample T-Test (One-Tailed) for the Mean Child Score/3 on the Model-Score-1 Trials – Ages 5-10 (Whole Years) in the Increased Information (I), Temporal (T), Spatial (S) and Temporal-Spatial (TS) Conditions

Age (whole years)	Condition	Model-score-1 trials			
		Child mean score/3		T-test – chance- level score (1.000)	
		Mean	SD	t	p
5	I	1.810	0.430	9.970	< .001
6	I	1.817	0.383	11.869	< .001
7	I	1.769	0.479	8.195	< .001
8	I	1.917	0.359	13.526	< .001
9	I	1.935	0.303	17.173	< .001
10	I	1.817	0.382	9.561	< .001
5	T	1.595	0.419	7.513	< .001
6	T	1.806	0.402	11.180	< .001
7	T	1.833	0.392	10.851	< .001
8	T	1.881	0.275	16.927	< .001
9	T	1.839	0.284	16.456	< .001
10	T	1.850	0.315	12.074	< .001

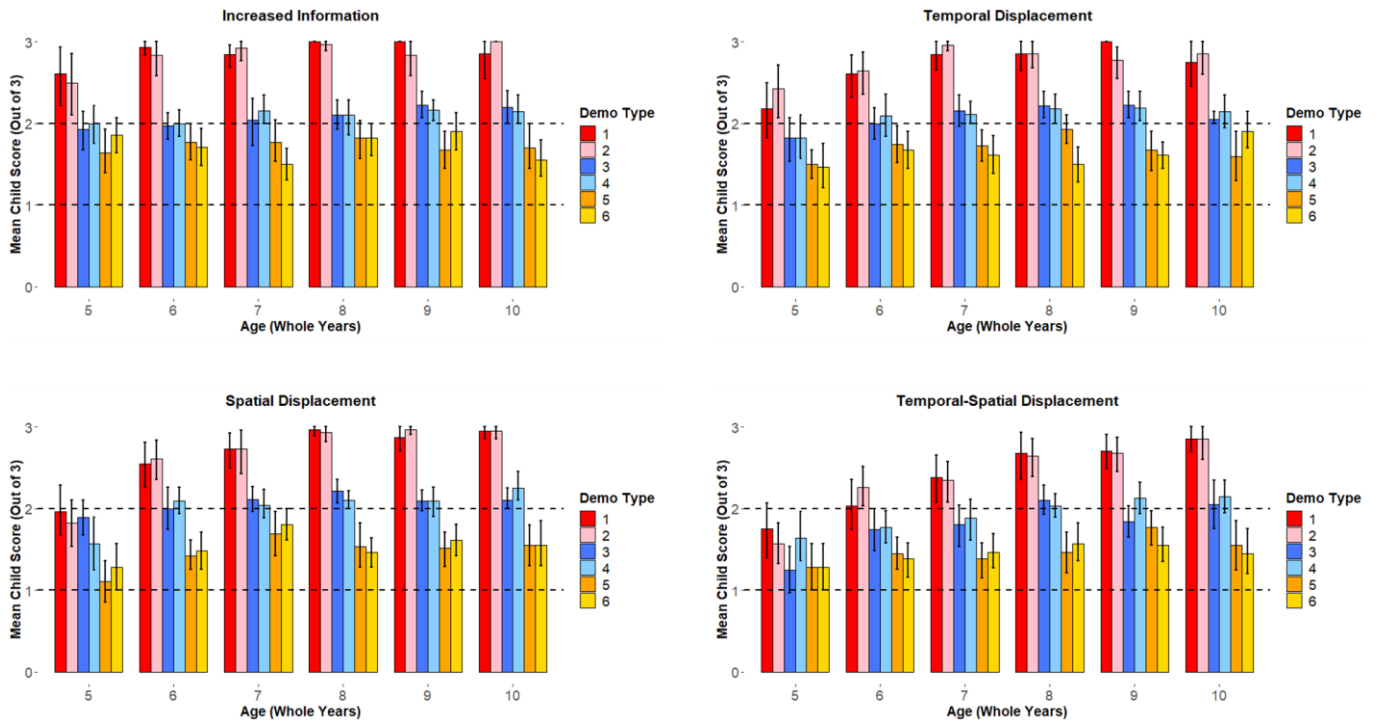
5	S	1.429	0.361	6.277	< .001
6	S	1.634	0.348	10.148	< .001
7	S	1.872	0.284	15.650	< .001
8	S	1.738	0.367	10.645	< .001
9	S	1.742	0.410	10.072	< .001
10	S	1.733	0.298	11.000	< .001
5	TS	1.274	0.506	2.866	0.004
6	TS	1.527	0.411	7.142	< .001
7	TS	1.551	0.421	6.677	< .001
8	TS	1.714	0.283	13.367	< .001
9	TS	1.720	0.323	12.413	< .001
10	TS	1.683	0.439	6.962	< .001

Note. The one-sample t-test was used to determine if the mean child score (out of 3) was significantly greater than a chance-level score (1).

Figures

Figure S.3

Mean Child Score/3 by Demonstration Type and Age (Whole Years) for the Increased Information, Temporal, Spatial and Temporal-Spatial Conditions



Note. The dashed line at 1 depicts chance performance: the score expected if children were selecting chests at random and not using the social information. The dashed line at 2 depicts the highest single model score on any trial.

Experimenter Verbal Scripts – Chapters 2-5

Verbal Script: Chapter 2

Enduring Cues Condition

Introduction

“We’re going to play a game. You are Pirate [child’s name] and this is Polly the Parrot. We are going to look for hidden treasure together and see who can find the most treasure. Look at these [red] chests. One of these has treasure inside it but the others just have this paper. How many [red] chests have treasure inside them?”

“In this game we have to mix the chests up so we don’t know where the treasure is.”
(Chests are placed into a sack and mixed up before the next two colours are given to the child to examine and place into the sack as above).

“Ok, now let’s collect lots of treasure with Polly the Parrot, arrrrr. Remember to try and find more treasure than Polly.”

Puppet Selections x 3 (only after all three puppet selections does the child make their selections).

([Red] chests are lifted out of the sack and placed onto the map).

Instructions

“Polly the Parrot’s turn to collect [red] treasure.”

Response (selected chests opened to reveal contents and then left partially open so contents visible)

Rewarded: “There it is! Polly the Parrot found where the [red] treasure is and can now put a piece of this treasure on her [red] square.” *(Polly squawks)*

Unrewarded: “Oh no, the treasure isn’t in that [red] chest.”
(Polly shakes head)

Child Selections x 3

Instructions

“When I say go, it is your turn to collect [red] treasure. You can open this one, this one or this one *(motion to the three [red] chests in turn – direction counterbalanced)* to try and get more treasure than Polly. Ok, go!”

Response

Rewarded: “There it is! That’s where the [red] treasure is! You can now put a piece of treasure on your [red] square.” *(Polly squawks)*

Unrewarded: “Oh no, the treasure isn’t in that [red] chest.”
(Polly shakes head)

Between Trials

“Do you want to look for some more treasure with Polly the Parrot? Do you think you can find more treasure than Polly?”

Transient Information Condition

Introduction

“We’re going to play a game. You are Pirate [child’s name] and this is Polly the Parrot. We are going to look for hidden treasure together and see who can find the most treasure. Look at these [red] chests. One of these has treasure inside it but the others just have this paper. How many [red] chests have treasure inside them?”

“In this game we have to mix the chests up so we don’t know where the treasure is.”

(Chests are placed into a sack and mixed up before the next two colours are given to the child to examine and place into the sack as above).

“Ok, now let’s collect lots of treasure with Polly the Parrot, arrrrr. Remember to try and find more treasure than Polly.”

Puppet Selections x 3 (only after all three puppet selections does the child make their selections).

([Red] chests are lifted out of the sack and placed onto the map).

Instructions

“Polly the Parrot’s turn to collect [red] treasure.”

Response (selected chests opened to reveal contents and then closed)

Rewarded: “There it is! Polly the Parrot found where the [red] treasure is and can now put a piece of this treasure on her [red] square.” *(Polly squawks)*

Unrewarded: “Oh no, the treasure isn’t in that [red] chest.”

(Polly shakes head)

Memory Questions x 3

(Ask these two questions right before the child makes their selection for each colour in an array).

Where did Polly the Parrot look for the [red] treasure?

Did Polly find the [red] treasure?

Child Selections x 3

Instructions

“When I say go, it is your turn to collect [red] treasure. You can open this one, this one or this one *(motion to the three red chests in turn – direction counterbalanced)* to try and get more treasure than Polly. Ok, go!”

Response

Rewarded: “There it is! That’s where the [red] treasure is! You can now put a piece of treasure on your [red] square.” (*Polly squawks*)

Unrewarded: “Oh no, the treasure isn’t in that [red] chest.”

(*Polly shakes head*)

Between Trials

“Do you want to look for some more treasure with Polly the Parrot? Do you think you can find more treasure than Polly?”

Verbal Script: Chapter 3

Introduction

“Hi [child’s name], do you want to search for treasure today?”

“Great!”

Instructions

“The treasure is gold coins, like these *(point to coin on screen)*. The treasure is hidden inside chests, like these *(point to grid of chests on screen)*. There is treasure inside three chests *(hold up 3 fingers)* and you should try to get all three pieces of treasure. If you find the treasure, you will see a coin inside the chest *(point to this on screen)*. If you don’t find the treasure, you will see a cross inside the chest *(point to this on screen)*. Pirate [octopus/parrot] *(point to image on screen)* and pirate [parrot/octopus] *(point to image on screen)* will also be looking for treasure, so you should watch them carefully.”

Increased Information Condition

Practise Trial

“Ok, so now we can practise together. Are you ready?”

“First, pirate [octopus/parrot] looks for the treasure. Then, pirate [parrot/octopus] looks for the treasure.”

“Now it’s your turn – remember, you should try to get all three pieces of treasure and you can choose a chest even if one of the animals has already chosen it.”

If the child fails to select at least 1 of the rewarded chests selected by the models say:

“Remember, you can choose a chest even if one of the animals has already chosen it.”

“Look – that’s where the treasure was *(point to the open chests)*. How many pieces of treasure were there?”

“Yes – 3 pieces. There will always be 3 pieces of treasure. Are you ready to start for real?”

“Great! Remember, pirate octopus and pirate parrot will go first and then you need to try and get all three pieces of treasure.”

6 trials

“Ready?”

“Pirate [octopus’s/parrot’s] turn.”

“Pirate [parrot’s/octopus’s] turn.”

“Your turn.”

“That’s where the treasure was.”

Temporal Displacement Condition**Practise trial**

“Ok, so now we can practise together. Are you ready?”

“First, pirate [octopus/parrot] looks for the treasure, and now the chests have closed.”

“Then, pirate [parrot/octopus] looks for the treasure, and now the chests have closed.

Now it’s your turn – remember, you should try to get all three pieces of treasure and you can choose a chest even if one of the animals has already chosen it.”

If the child fails to select at least 1 of the rewarded chests selected by the models say:

“Remember, you can choose a chest even if one of the animals has already chosen it.”

“Look – that’s where the treasure was (*point to the open chests*). How many pieces of treasure were there?”

“Yes – 3 pieces. There will always be 3 pieces of treasure. Are you ready to start for real?”

“Great! Remember, pirate octopus and pirate parrot will go first and then you need to try and get all three pieces of treasure.”

6 trials

“Ready?”

“Pirate [octopus’s/parrot’s] turn.”

“Pirate [parrot’s/octopus’s] turn.”

“Your turn.”

“That’s where the treasure was.”

Spatial Displacement Condition

Practise trial

“Ok, so now we can practise together.”

“Are you ready?”

“This time there are three sets of chests – look *(point to each set in turn)*, and the treasure is in the same place on each set. First, pirate [octopus/parrot] looks for the treasure in this set of chests.”

(point to top left set)

“Then, pirate [parrot/octopus] looks for the treasure in this set of chests *(point to top right set)*. Now, it’s your turn to look for the treasure in this set of chests *(point to bottom set)* – remember, you should try to get all three pieces of treasure and you can choose a chest even if one of the animals has already chosen it.”

If the child fails to select at least 1 of the rewarded chests selected by the models say:

“Remember, you can choose a chest even if one of the animals has already chosen it.”

“Look – that’s where the treasure was *(point to open chests in child’s set)*, it’s in the same place on each set of chests *(point to 3 sets of chests)*. How many pieces of treasure were there?”

“Yes – 3 pieces. There will always be 3 pieces of treasure. Are you ready to start for real?”

“Great! Remember, pirate octopus and pirate parrot will go first and then you need to try and get all three pieces of treasure.”

6 trials

“Ready?”

“Pirate [octopus’s/parrot’s] turn on this set of chests.”

“Pirate [parrot’s/octopus’s] turn on this set of chests.”

“Your turn on this set of chests.”

“That’s where the treasure was.”

Temporal-Spatial Displacement Condition

Practise trial

“Ok, so now we can practise together. Are you ready?”

“This time there are three sets of chests – look (*point to each set in turn*), and the treasure is in the same place on each set. First, pirate [octopus/parrot] looks for the treasure in this set of chests (*point to top left set*), and now the chests have closed.”

“Then, pirate [parrot/octopus] looks for the treasure in this set of chests (*point to top set*), and now the chests have closed. Now, it’s your turn to look for the treasure in this set of chests (*point to bottom set*) – remember, you should try to get all three pieces of treasure and you can choose a chest even if one of the animals has already chosen it.”

If the child fails to select at least 1 of the rewarded chests selected by the models say:

“Remember, you can choose a chest even if one of the animals has already chosen it.”

“Look – that’s where the treasure was (*point to open chests in child’s set*), it’s in the same place on each set of chests (*point to 3 sets of chests*). How many pieces of treasure were there?”

“Yes – 3 pieces. There will always be 3 pieces of treasure. Are you ready to start for real?”

“Great! Remember, pirate octopus and pirate parrot will go first and then you need to try and get all three pieces of treasure.”

6 trials

“Ready?”

“Pirate [octopus’s/ parrot’s] turn on this set of chests.”

“Pirate [parrot’s/ octopus’s] turn on this set of chests.”

“Your turn on this set of chests.”

“That’s where the treasure was.”

Ending

“Now let’s find out how many pieces of treasure you got..”

“You got [number of pieces of treasure displayed on screen] pieces of treasure!

Great treasure-hunting!!”

“You can choose a sticker. Which colour would you like?”

Verbal Script: Chapter 4

Part 1:

When Approaching Participant

Hen: "Hello, my name is Hen. What is your name?"

Hen: "Please can you help me find my chicks?"

In Testing Room

Hen: "I had some eggs with chicks in like these (*Show example.*) and some eggs with yolks in like these (*Show example.*) for Mr Fox. He can eat these. I put the eggs with chicks inside in one of these baskets, and the eggs that Mr Fox can eat in the other basket, but Mr Fox has mixed the baskets up and taken them both so I don't know which basket my chicks are in!"

Mr Fox: "You need to pay me coins if you want to look inside the eggs!"

Hen: "Oh no, I don't have any coins. I need you and my animal friends to help me find out which basket has my chicks inside". *Hen makes non-verbal actions towards the baskets to indicate she does not know where the chicks are.*

Hen: "Can you help?"

Hen: "Great. My animal friends will go first!"

Part 2: Condition 1 (uninformed demonstrator 1st and uninformed memory question 1st)

Uninformed 1: "*(Animal noise.)* My name is [Pig] and I can help! I have one coin Mr Fox, and I will buy one egg from you. Here is my coin Mr Fox".

Uninformed 1: "Hmmm I am not sure which basket the chicks are inside."

(Makes non-verbal actions to show uncertainty.)

Uninformed 1: "I choose this one!"

(Select incorrect basket.)

Experimenter: "I wonder if [Pig] has chosen the basket with chicks inside."

(Egg is placed onto the board.)

Experimenter: “[Pig] now needs to go and [pay in the mud].”

(Leaves and is hidden away.)

(Animal noise.)

Uninformed 2: “*(Animal noise.)* My name is [Cow] and I can help! I have one coin Mr Fox, and I will buy one egg from you. Here is my coin Mr Fox”.

Uninformed 2: “Hmmm I am not sure which basket the chicks are inside.”

(Makes non-verbal actions to show uncertainty.)

Uninformed 2: “I choose this one!”

(Select incorrect basket.)

Experimenter: “I wonder if [Cow] has chosen the basket with chicks inside.”

(Egg is placed onto the board.)

Experimenter: “[Cow] now needs to go and [eat some grass].”

(Leaves and is hidden away.)

(Animal noise.)

Uninformed 3: “*(Animal noise.)* My name is [Duck] and I can help! I have one coin Mr Fox, and I will buy one egg from you. Here is my coin Mr Fox”.

Uninformed 3: “Hmmm I am not sure which basket the chicks are inside.”

(Makes non-verbal actions to show uncertainty.)

Uninformed 3: “I choose this one!”

(Select incorrect basket.)

Experimenter: “I wonder if [Duck] has chosen the basket with chicks inside.”

(Egg is placed onto the board.)

Experimenter: “[Duck] now needs to go and [have a bath].”

(Leaves and is hidden away.)

(Animal noise.)

Informed (1st choice): “*(Animal noise.)* My name is [Dog] and I can help! I have three coins Mr Fox, and I will buy one egg from you now *(Picks up 1st coin.)*. Here is my first coin Mr Fox. Hmmm I am not sure which basket the chicks are inside *(Makes non-verbal actions to show uncertainty.)*. I choose this one *(Select incorrect basket.)* but I better have a look inside before I make my next choice.”

(Open egg slightly and hold to the animal's eye.)

Experimenter: "I wonder if [Dog] has chosen the basket with chicks inside."

(Egg is placed onto the board.)

Informed (2nd choice): *(Picks up 2nd coin)* "Here is my second coin Mr Fox. I would like to buy another egg from you Mr Fox. This time I choose this basket and will look inside my egg."

(Select correct basket and look inside egg.)

Experimenter: "I wonder if [Dog] has chosen the basket with chicks inside."

(Egg is placed onto the board.)

Informed (3rd choice): *(Picks up 3rd coin)* "I can buy one more egg from you Mr Fox. Here is my third coin".

(Select correct basket and look inside egg.)

(Egg is placed onto the board.)

Experimenter: "[Dog] now needs to go and [play with a ball]."

(Animal noise.)

(Informed demonstrator is hidden away.)

Child's Choice (all animals should be hidden away)

Experimenter: "There are still lots of eggs with chicks inside that need to be rescued. I tell you what, I have got many coins (*Bring out chest with extra coins inside and show child.*) and we can buy ONE basket from Fox with those coins. Which basket should we buy to rescue the chicks? Point to it."

(Record response.)

Experimenter: "Is that OK Mr Fox?"

Mr Fox: Yes.

Experimenter: “Thank you Mr Fox. We will come back and take one of the baskets in a minute.”

Remaining Questions and Script (*NOTE – the order of these will differ between participants)

Experimenter: “Why did you choose that basket?” *(Record child’s response.)*

Experimenter: “Which basket did *Uninformed* (actual animal differs between participants) look inside?” *(Record child’s response.)*

Experimenter: “Which basket did *Informed* (actual animal differs as above) look inside on their last turn?” *(Record child’s response.)*

Experimenter: “Which animal had the most coins?” *(Record child’s response.)*

Experimenter: “You can now take the basket you pointed to.”

(Experimenter to pass the selected basket to the child and prompt them to open an egg to see if it contains a chick.)

Experimenter: “Well done you found Hen’s chicks!”

(Sit Hen on top of the correct basket. Child can also be allowed to peak into an egg from the other basket if they wish.)

OR

Experimenter: “Hen’s chicks weren’t in that basket, were they? Maybe if you ask nicely Mr Fox will let you swap the baskets so that you can save the chicks. Can you ask Mr Fox?”

(Child to ask Mr Fox if this is ok and he says yes!)

Experimenter: “Well done you found Hen’s chicks!”

(Sit Hen on top of the correct basket.)

(Child is then given a sticker to thank them for their participation (except in University Kindergarten)).

Part 2: Condition 2 (script is the same other than the choices of the informed demonstrator – see below)

Informed (1st choice): “*(Animal noise)* My name is [Dog] and I can help! I have three coins Mr Fox, and I will buy one egg from you now *(Picks up 1st coin)*. Here is my first coin Mr Fox. Hmm I am not sure which basket the chicks are inside.”

(Makes non-verbal actions to show uncertainty.)

Informed (1st choice): “I choose this one *(select correct basket)* but I better have a look inside before I make my next choice.”

(Open egg slightly and hold to animal’s eye.)

Experimenter: “I wonder if [Dog] has chosen the basket with chicks inside.”

Egg is placed onto the board

Informed (2nd choice): *(Picks up 2nd coin)* “Here is my second coin Mr Fox. I would like to buy another egg from you Mr Fox”.

(Select correct basket and look inside egg.)

Experimenter: “I wonder if [Dog] has chosen the basket with chicks inside.”

(Egg is placed onto the board.)

Informed (3rd choice): *(Picks up 3rd coin.)* “I can buy one more egg from you Mr Fox. Here is my third coin”.

(Select correct basket and look inside egg.)

(Egg is placed onto the board.)

Experimenter: “..... now needs to go and.....”

(Animal noise.)

(Informed demonstrator is hidden away.)

Verbal Script: Chapter 5

Introduction Trial 1

Approaching the Participant

Hen: "Hello, my name is Hen. What is your name?"

Testing Room

Hen: "I have something to show you. Look at these eggs (*show 2 example eggs – closed*). They look the same on the outside, but you can open them and see what is inside (*child prompted to open eggs – help given if needed*). This egg (*point at contents of egg with chick inside*) has one of my baby chicks inside. Can you see my baby chick? This other egg (*point at contents of the other egg*) is for eating. See? The eggs look the same on the outside, but some have my baby chicks inside and some do not. I really need your help! I have done something silly and lost my baby chicks! I have four yellow baskets and my chicks are inside one of them, but I don't know which basket I put the chicks inside. Can you help me find out which basket has my baby chicks inside?"

"Great, thank you!"

OR

"That's OK – would you like a sticker and we can go back to class?"

"I am going to go away now and try to find some other animals to help find my baby chicks. I am a very silly hen! A very silly hen indeed.... (*Hen exits the scene muttering that she is a silly hen...*)

Demonstrations Trial 1

Uninformed demonstrator (although half the time informed will go 1st):

Uninformed demonstrator: "*(Animal noise)* Hello, my name is and I will try to help you find the basket with the chicks inside! Hmmm, but I'm not sure which basket the chicks are inside." (*Animal shakes head from side to side.*)
(Hovers over the 3 baskets they will NOT choose and stops to look inside each, sticking head inside slightly to indicate looking. The basket the puppet finally selects

should be looked in last.)

Experimenter: “Look at what is doing!” *(This is said as the animal begins to look inside the baskets.)*

Then the animal hovers over the basket they will choose (a pre-determined unrewarded basket in this case), looks inside as above and immediately states:

Uninformed demonstrator: “I choose this one! Hen should keep this basket.”

(Animal grabs hold of the basket and picks it up/shakes it slightly, making noise to aid the child’s memory of the selection.)

Experimenter: “I wonder if has chosen the basket with the baby chicks inside.”

Uninformed demonstrator: “OK, I need to go home now! Bye!”

(Animal exits the scene whilst making their noise.)

Informed demonstrator

Informed demonstrator: “*(Animal noise.)* Hello, my name is and I will try to help you find the basket with the chicks inside! Hmmm, but I’m not sure which basket the chicks are inside.” *(Animal shakes head from side to side.)*

(Hovers over the 3 baskets they will NOT choose and stops to look inside each, sticking head inside slightly to indicate looking AND picking up, and peeking inside, one egg from each basket. The basket the puppet finally selects should be looked in last.)

Experimenter: “Look at what is doing!” *(This is said as the animal begins to look inside the baskets.)*

Then hovers over the basket they will choose (a pre-determined rewarded basket in this case), looks inside and peeks into an egg as above and immediately states:

Informed demonstrator: “I choose this one! Hen should keep this basket.”

(Animal grabs hold of the basket and picks it up/shakes it slightly, making noise to aid the child’s memory of the selection.)

Experimenter: “I wonder if has chosen the basket with the baby chicks inside.”

Informed demonstrator: “OK, I need to go home now! Bye!”

(Animal exits the scene whilst making their noise.)

Child’s Choice Trial 1

Experimenter: “Now you can choose which basket you think we should give to Hen. You should choose the basket that you think has Hen’s baby chicks inside because Hen is looking for her baby chicks.”

Q1: “Which basket should we give to Hen? Point to it.” *(Record response.)*

Further Questions Trial 1

Experimenter: “You chose this basket” *(motion to the chosen basket).*

Q2: “Why did you choose this basket?” *(Record response.)*

Q3: “What do you think is inside this basket?” *(Record response – only ask if the child does not include this information in their answer to the above question.)*

Q4: “Which basket did Cow choose?” *(Record response.)*

Q5: “Which basket did Dog choose?” *(Record response.)*

Experimenter to child: “OK, we can now put this basket over here for Hen.” *(Basket placed out of reach but in sight.)*

Getting ready for trial 2

The experimenter chats to the child about their favourite animals.

Charlotte to quickly put the remaining baskets/eggs from trial 1 into a box (hidden under table) and lift out the 4 green baskets for trial 2 (which will be ready and in the correct order under the table).

Introduction Trial 2

Hen: “Hello, I really need your help again! I have done something silly and lost more of my baby chicks! I have four green baskets and my chicks are inside one of them, but I don’t know which basket I put the chicks inside...”

Experimenter: “Silly Hen has lost her chicks *again!* Hen really is *very* silly!”

Hen: “Can you help me find out which basket has my baby chicks inside?”

Hen: “Great, thank you!”

OR

“That’s OK – would you like a sticker and we can go back to class?”

Hen: “I am going to go away now and try to find some other animals to help find my baby chicks. I am a very silly hen! A very silly hen indeed.... *(Hen exits the scene muttering that she is a silly hen...)*

Demonstrations Trial 2

Informed demonstrator (opposite demonstrator to trial 1 goes 1st)

Informed demonstrator: “*(Animal noise.)* Hello, my name is and I will try to help you find the basket with the chicks inside! Hmmm, but I’m not sure which basket the chicks are inside.” *(Animal shakes head from side to side.)*

Hovers over the 3 baskets they will NOT choose and stops to look inside each, sticking head inside slightly to indicate looking AND picking up, and peeking inside, one egg from each basket. The basket the puppet finally selects should be looked in last.

Experimenter: “Look at what is doing!” *(This is said as the animal begins to look inside the baskets.)*

Then hovers over the basket they will choose (a pre-determined rewarded basket in this case), looks inside and peeks into an egg as above and immediately states:

Informed demonstrator: “I choose this one! Hen should keep this basket.”

(Animal grabs hold of the basket and picks it up/shakes it slightly, making noise to aid the child’s memory of the selection.)

Experimenter to child: “I wonder if has chosen the basket with the baby chicks inside.”

Informed demonstrator: “OK, I need to go home now! Bye!”

(Animal exits the scene whilst making their noise.)

Uninformed demonstrator

Uninformed demonstrator: “*(Animal noise.)* “Hello, my name is and I will try to help you find the basket with the chicks inside! Hmmm, but I’m not sure which basket

the chicks are inside.” *(Animal shakes head from side to side.)*

Hovers over the 3 baskets they will NOT choose and stops to look inside each, sticking head inside slightly to indicate looking. The basket the puppet finally selects should be looked in last.

Experimenter: “Look at what is doing!” *(This is said as the animal begins to look inside the baskets.)*

Then the animal hovers over the basket they will choose (a pre-determined unrewarded basket in this case), looks inside as above and immediately states:

Uninformed demonstrator: “I choose this one! Hen should keep this basket.”

(Animal grabs hold of the basket and picks it up/shakes it slightly, making noise to aid the child’s memory of the selection.)

Experimenter: “I wonder if has chosen the basket with the baby chicks inside.”

Uninformed demonstrator: “OK, I need to go home now! Bye!”

(Animal exits the scene whilst making their noise.)

Child’s Choice Trial 2

Experimenter: “Now you can choose which basket you think we should give to Hen. You should choose the basket that you think has Hen’s baby chicks inside because Hen is looking for her baby chicks.”

Q1: “Which basket should we give to Hen? Point to it.” *(Record response.)*

Further questions trial 2

Experimenter: “You chose this basket” *(motion to the chosen basket).*

Q2: “Why did you choose this basket?” *(Record response.)*

Q3: “What do you think is inside this basket?” *(Record response – only ask if the child does not include this information in their answer to the above question.)*

Q4: “Which basket did Pig choose?” *(Record response.)*

Q5: “Which basket did Duck choose?” *(Record response.)*

BELOW QUESTION ONLY ASKED IN TRIAL 2 (this trial)!

Q6: Experimenter: “Did any of the animals (*hold up both puppets*) help you decide which basket to choose?” (*Record response.*)

Experimenter: “OK, we can now put this basket over here for Hen.” (*Basket placed out of reach but in sight.*)

Ending

Experimenter: “OK, so now we can look to see if you have found Hen’s chicks”.
(*Experimenter to pass the chosen baskets from trial 1 and trial 2 to the child and prompt them to open an egg from each to see if it contains a chick.*)

(*Hen reappears.*)

Hen to child: “Wow, you have found all/lots of my baby chicks. Thank you so much for your help! You have made me a very happy hen.”

OR

Hen to child: “My chicks are not in those baskets. But do not worry, I will carry on looking for my chicks. Thank you so much for your help! You have made me a very happy hen.”

The experimenter then offers the child a sticker as a thank you for their participation (except in University Kindergarten).