UNIVERSITY OF SOUTHAMPTON

FACULTY OF HUMANITIES

Archaeology

Lithic Morphological Variability as a Proxy for Palaeolithic Linguistic Ability: A Knapping Training Study Exploring Cultural Transmission, Theory of Mind and Language

by

Cory Marie Stade

Thesis for the degree of Doctor of Philosophy

June 2017

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF HUMANITIES

Archaeology

Doctor of Philosophy

Lithic Morphological Variability as a Proxy for Palaeolithic Linguistic Ability:

A Knapping Training Study Exploring Cultural Transmission, Theory of Mind and Language

by Cory Marie Stade

The identification of Palaeolithic symbolic material culture has often been taken as an indication of cognitive complexity, which limits interpretations to rare, late and localised events. This thesis develops an alternative method of interpretation through an empirical and psychologically supported chain of inference which avoids these limitations. It tests a hypothesis which predicts that material culture produced by varying methods of social learning will differ in their range of relative morphological variability due to varying levels of copy error. For Palaeolithic assemblages, high levels of standardisation might only be possible with the high fidelity transmission involved in complex social learning methods which utilise theory of mind (the ability to think about thoughts). Theory of mind, in turn, is highly correlated to linguistic ability; as both these abilities are gradient, using theory of mind as a proxy for language ability means Palaeolithic communities can be attributed with not just a binary presence or absence of language, but semantic and syntactic abilities which correlate with stages of theory of mind acquisition in modern humans. This thesis presents a knapping experiment where participants replicate model handaxes in different simulated social learning environments. Results show that different levels of morphological variation in the tools they produce support that high fidelity social learning methods produce a lower range of morphological variability. Results suggest that the origins of language lie somewhere before the origins of Acheulean technology, and that teaching and grammatical language were in the behavioural repertoires of Middle Palaeolithic hominins.

List of Contents

List of Figures

List of Tables

Academic Thesis: Declaration Of Authorship

I, ………………………………………………………………………………………… [please print name]

declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

[title of thesis] ……………………………………………………………………………………………………..

……

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Either none of this work has been published before submission, or parts of this work have been published as: [please list references below]:

Signed: ………………………………………………………………………

Date: ………………………………………………………………………

Acknowledgements

This thesis is no different than any other, in that it is a product of cumulative cultural knowledge. It has been gratefully culturally transmitted to me by a number of individuals I am forever indebted to. CAHO is a stimulating and exciting Palaeolithic research group to be a part of, and the John Wymer Lab, surrounded by the Gamble Library, was an inspirational place filled with things and people to 'think through'. Clive Gamble, ever capable of inspiring and encouraging one to 'think big', was a source of intellectual ideas with a topic that could move in so many directions. John McNabb was expertly able to ask the right questions always, and demand clarity and rigorous explanations of my choices at every stage. Rebecca Farbstein, along with Mac, provided extremely useful feedback at my PhD upgrade and put me on a steady track toward the finish line. I thank Christian Hoggard for introducing me to the world of geometric morphometrics and keeping me motivated through the final sprint to the end. Patrick Cuthbertson made the experiment at all possible, and the carrying out of the logistics was completely a joint effort. Friends were ever kind in providing help (or a ride in their car). I thank Klint Janulis, Amanda Janulis, Ringo Janulis, Manse Ahmad, Lee Graña, Corie Graña, and Anna-Elise Young for their support. A number of people read parts of this thesis (in addition to those listed above) and offered advice, or contributed to my ideas through a useful chat about topics: Mark Collard, Frederick Coolidge, James Dilley, Huw Groucutt, Brian Hayden, Cecilia Heyes, Sverker Johansson, Nada Khreisheh, Bethan Linscott, Annemieke Milks, Iza Romanowska, Eleanor Scerri, Ceri Shipton, and Thomas Wynn all helped me refine my arguments. I thank Bath Potters for their porcelain clay and careful firing of the experiment material, and Fusion Arts Centre, Oxford for providing the perfect indoor space for the knapping experiment to be carried out. Finally I thank all the participants who took part in my pilot and my main experiment – your enthusiasm and fantastic porcelain handaxes impressed and pleased me in equal measures.

Glossary of Terms

Complementation: a simple grammatical construction where one linguistic element requires another to complete its meaning.

Copy error: Mistakes in replication

Cultural transmission: The exchange of information in a non-genetic manner

Emulation: End-state copying behaviour; reproducing an item without regard to the processes that produced it

Imitation: Process-based copying behaviour; reproducing an item by intentionally copying the process that produced it

Intentionality: "aboutness"; a way to describe the embedded levels of theory of mind operations (second level intentionality is basic theory of mind).

Language: A learned system of symbols for intentional communication

Lexicon: A repertoire of words in a language

Mimicry: Copying an action without regard for its purpose (for example, simply as a social engagement)

Semantics: the relationship between symbols and what they denote; meaning

Social learning: The acquisition of new behaviours by cultural transmission

Symbol: An arbitrary conceptual unit with a learned, communally agreed referent

Syntax: A system of rules for organizing linguistic elements into larger units to create new meanings

Teaching: Intentionally conferring knowledge to another

Theory of mind (ToM): The ability for an individual to theorise about the mental states of others

Chapter One.

Cognitive Fossils

1.1 Introduction

This thesis is about the archaeology of language. A "wry quip often made" (Botha and Everaert 2013, p. 1) regarding language is that it does not fossilise, in that it does not leave direct traces in the material record, especially before the advent of writing systems or the appearance of referential art:

"Fading as soon as it is uttered, spoken language leaves no trace." (Burling 2005, p. 1)

"…language has left no obvious equivalent to fossil teeth and bones, and seems inaccessible to enquiry." (Solé 2005, p. 289)

"Behavior, famously and axiomatically, does not fossilize, and communication events are the most ephemeral of behaviors…" (Armstrong and Wilcox 2007, p. 9)

"…language does not fossilize, and we have no time machines." (Fitch 2010, p. 6)

"Language does not fossilize like bones and we have no "fossils" that would help us to reconstruct the phylogeny of language." (Kováč 2015, p. 58)

But while language does not experience a process of petrification, as bone or other once-organic specimens might, this thesis makes the claim that language ability is detectable in stone tool technology; archaeologists can detect complex social learning (the method by which the knowledge was culturally transmitted), the associated level of theory of mind ability required (the perceptual information acquired by theorizing about the intentions and knowledge of others), and the associated linguistic skills required to develop those theory of mind abilities (such as words and grammar which lead species to appreciate the existence of other minds).

This thesis approaches the mind as embodied (where the physical world shapes cognitive processes and can therefore not be excluded from their characterisation, Malafouris 2013), as well as extended and distributed (where cognitive processes are not 'all in the head', but extend beyond the confines of the skull and even the body, Clark and Chalmers 1998). This theoretical position is important for reasons of legitimacy in reconstructing hominin minds from materials they created. It is essentially a materialist standpoint, where material is foundational to all things. As this thesis concerns connecting materials to minds, this materialist approach makes the connection possible. In an extended, distributed mind, objects in the environment (including stone tools) function as part of the mind's processes. In an embodied mind, tools are therefore not just objects that intentional agents act upon, but agents themselves and *things we think through* (Malafouris 2013). Palaeolithic stone tools can therefore be seen as 'cognitive fossils': in the way a bone was once part of a working skeleton, a handaxe was once part of a mental act.

From this, researchers can look to how stone tools vary as a result of behaviour, and make cognitive inferences about the differing cognitive acts the behaviours may have been involved with. A multitude of factors impact how similar stone tools are to each other within and between assemblages (for instance, variation could be introduced because of raw material, function or skill, among other reasons), and this thesis explores the impact of teaching and learning on lithic morphological variability. Mesoudi and O'Brien (2008, p. 3) suggested that differences in transmission style may be detectable in the archaeological record, as changes in social learning would have significant effects on communities that might manifest in the material. Eerkens and Lipo (2005) have also suggested information concerning social learning could be detectable archaeologically, and that changes in morphological variation could identify different types of social learning in archaeological material. However, Tehrani and Riede (2008, p. 235) note that archaeologists "have rarely sought explicitly to discriminate between specific modes of cultural transmission" in the way material technologies are proposed to have been learned (but see recent work by Schillinger et al. 2015, Morgan et al. 2015 for examples of research that deals with these questions). This highlights the importance and necessity of inquiries such as in this thesis.

The central question this thesis addresses is whether or not there is variation that can be explained as a result of social learning processes, using a controlled experiment and the data created therein, and whether that information can then be extended to provide information about theory of mind and linguistic ability among Palaeolithic populations.

1.2 The Research Question

This thesis will use the existing literature (described further in Section 1.5) to construct an argument that language ability and theory of mind ability have a strong correlative relationship such that a co-evolution of the two was likely. This correlative relationship would have probably existed in ancestral hominins, such that theory of mind level would indicate linguistic level in prehistory as well as in modern populations. As the literature can be interpreted to suggest that theory of mind enables certain modes of social learning, and that social learning impacts copy error in material culture, a chain of inference from social learning method to required theory of mind level to required linguistic level will be formed (Figure 1.1). The research question for this thesis is, then,

Does social learning influence the range of morphological variability of lithic assemblages, and if so can this effect be used to deduce social learning method, theory of mind ability, and therefore language ability amongst Palaeolithic stone tool makers?

The research question is what guides the formation of the hypothesis and ultimately this thesis' experiment.

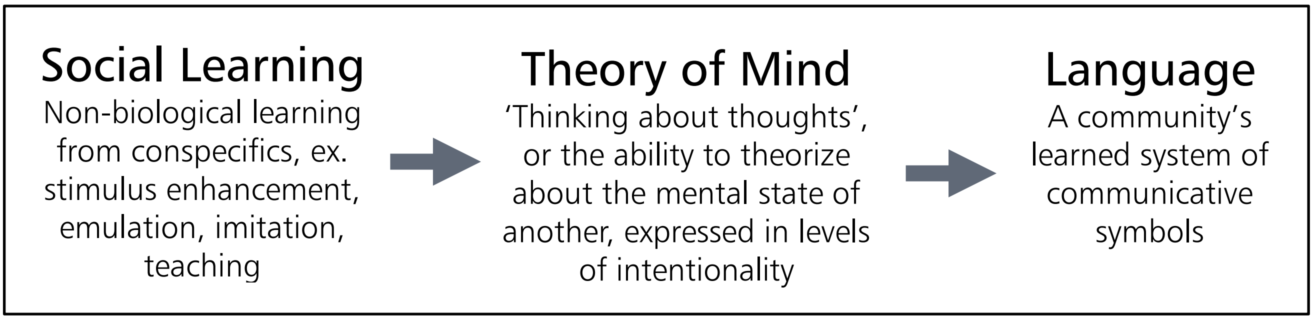


Figure 1.1 A chain of interpretive reasoning: the mode of cultural transmission indicates necessary theory of mind ability, which requires certain linguistic abilities for theory of mind's development.

1.3 The Hypothesis

The hypothesis that underpins this thesis is that variation in lithic assemblages are affected by the degree of theory of mind used in its social transmission, as it gives learners differential access to knowledge (Figure 1.2). Different ranges of morphological variation in replicated materials are hypothesised because the fidelity of the knowledge transfer will affect the rate of copy error within the transmission of tool-making knowledge (Eerkens and Lipo 2007). High fidelity social learning methods which utilise theory of mind (imitation and teaching) have a higher faithfulness of information transfer (lower copy error, due to psychophysical limits of human perception affecting the ability to accurately replicate them, Eerkens and Lipo 2005), and will therefore produce assemblages with less variability (higher standardisation) than social learning methods with a lower theory of mind requirement and lower fidelity of information transfer. Building from these studies, technical knowledge learned by social learning methods that require more complex theory of mind abilities and result in less copy error would hypothetically reduce tool variability, as items would be created more faithfully to the way they were taught, and therefore to others taught in the community. The highest levels of standardisation in an assemblage might therefore only be possible where teaching is the method of cultural transmission.

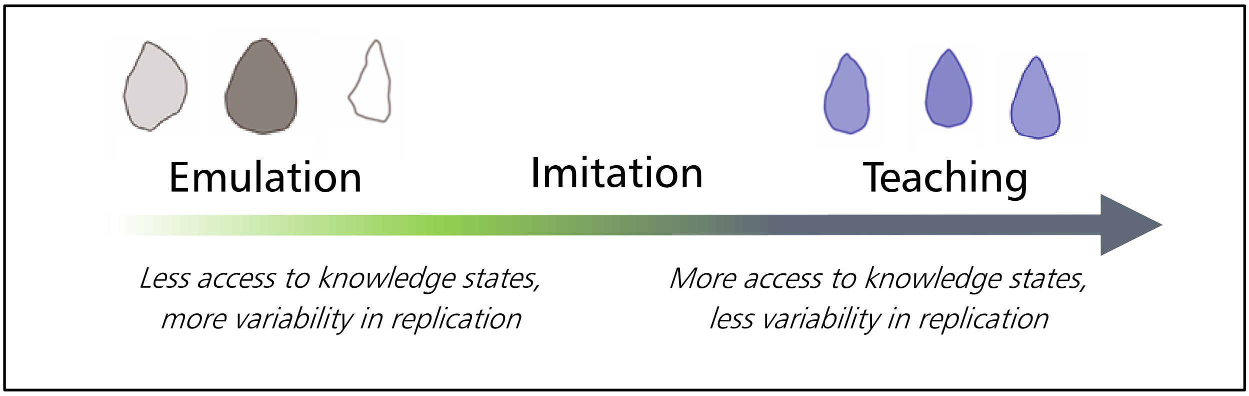


Figure 1.2 Three social learning methods arranged by their hypothesised effect on the variability of a lithic assemblage. The arrow represents increased theory of mind ability.

To reiterate, because of the different levels of theory of mind involved in different modes of cultural transmission, a chain of inference from cultural transmission ability, to theory of mind ability, to linguistic ability has been developed (Figure 1.1). Archaeological evidence of activities indicating specific social learning methods will then indicate the prerequisite theory of mind abilities for their transmission, in turn suggesting correlated linguistic abilities.

This thesis therefore addresses the key aim in evolutionary cognitive archaeology: to develop a method that assesses the cognitive abilities of Palaeolithic hominins, and specifically their language abilities. This study provides necessary empirical support for an often suggested link between cognition and tool standardisation, a claim repeatedly made in Palaeolithic research (for example in terms of a 'mental template', Mellars 1996, Monnier 2006, Chase 2008), but which requires explicit experimental support to exist as more than a common assumption. Uniquely, this thesis will use experimental archaeological evidence to support the presence of specific linguistic features, rather than the mere presence or absence of language. For example, imitation will be argued to denote the presence of a lexicon (a mental repertoire of words), as child development studies demonstrate how the acquisition of theory of mind, essential for imitation, is built upon pre-adaptations like joint attention, pointing and gaze following, and triangulation between self, object and other (Miller 2006). These skills are what make word reference and symbolic reference possible, and are both driven by and reinforce the acquisition of understanding others' intentions. Also from research in child development (de Villiers and Pyers 2002), teaching will be argued to evidence 1) the vocabulary to describe mental states, and 2) complementation (where one linguistic unit governs another in order to create its meaning, Crystal 1995). These skills are shown in child development studies to be key to the acquisition of third-level intentionality, which teaching relies on.

By using theory of mind as a proxy for cognition and language, and interpreting the complexity of language by inferred theory of mind abilities, this thesis represents a unique (and presently the sole) method of empirically supporting the presence of specific features of language ability in the archaeological record.

1.4 Thesis Rationale

A symbol is an arbitrary, conventionalised meaning that stands for something else (de Saussure et al. 2016); this Saussurian concept of 'arbitrariness of the sign' is a key characteristic of language, and is the nature of the relationship between symbols and referents in a linguistic community. In the archaeological record, beads and pendants are commonly taken to communicate the identity of an individual or group ethnographically, showing self and/or group awareness (Bar Yosef 2002). When archaeologists refer to populations as participating in symbolic behaviour by referencing symbolic material culture, or having "the ability to represent objects, people, and abstract concepts with arbitrary symbols, vocal or visual, and to reify such symbols in cultural practice" (McBrearty and Brooks 2000, p. 492), oftentimes language is interpreted as part of the 'symbolic package', the medium that allows the social landscape to be shared (Henshilwood and Marean 2003). Bar-Yosef (2002) identifies marine shells, teeth, ivory and eggshell (when interpreted as body decorations), as "clear signs for the identity of the social units" in past human societies. Use of symbols are also interpreted as showing the ability to utilise abstract concepts and the ability to 'outsource' the transmission of information to things other than themselves (Langley 2015).

Because of this commonality between attributing an artefact with a symbolic meaning, and assigning an arbitrary meaning to a communicative symbol through language, researchers often attribute the users of symbolic material culture with the ability to also use language, among a suite of other modern cognitive abilities:

"The only direct evidence for the first use of symbolic language amongst humans is the recognition in the archaeological record of the material products of symbolic thinking." (d'Errico et al. 2003, p. 6)

This type of interpretation highlights the structure of an interpretive model used in what I have termed 'the symbolic method':

- Symbolic behaviour can successfully be interpreted from material artefacts (often by a perceived lack of utilitarian use).
- This symbolic ability points to a rich symbolic life that includes cognitive abilities such as language.

Yet, few researchers actually break down this 'symbolism as proxy' argument into its constituent parts for analysis. Some criticisms of interpreting symbolism from material culture, in the case of shell beads, are discussed in Botha (2008, 2010; see Figure 1.3), and my own criticisms will be outlined here. For the purposes of utilising symbolism as a proxy for Palaeolithic cognition, the symbolic method suffers from four main issues:

1) A lack of explicit mechanisms detailing the cognitive requirements for symbolic thought and how these relate to language use;
2) The difficulty of demonstrating that an item was indeed symbolic to a Palaeolithic community, and what requirements an object must fulfil to be deemed symbolic by archaeologists;
3) Pre-Upper Palaeolithic symbolic objects are rare, and cognitive interpretations from symbolic material cannot be made in geographic areas or timescales where symbolic material is not in use or has not preserved, and the proxy is therefore limited to recent and localised spaces
4) The attribution of symbolism to an object allows only a 'presence or absence' interpretation, which results in a simple binary cognitive or

linguistic interpretation which does not reflect the scaled complexity of cognitive abilities such as language.

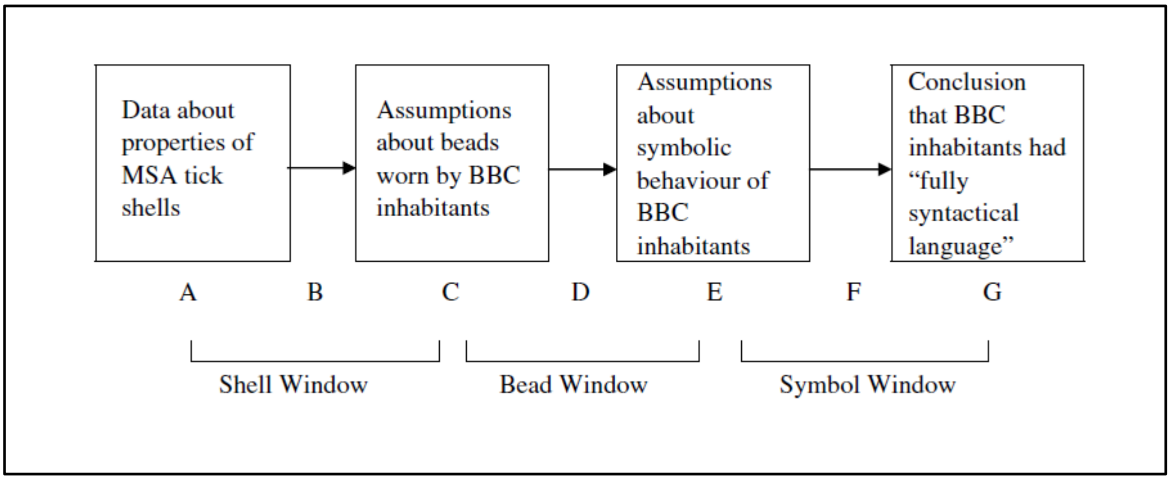


Figure 1.3. An image from Botha (2008) describing how Middle Stone Age pierced shells from Blombos Cave (BBC) are interpreted by some researchers as being symbolic and therefore provide evidence of syntactic language.

Theory of mind as an alternative proxy, however:

1) Has a well-defined predictive relationship with language ability (e.g. Miller 2006);
2) Has levels of complexity that correlate with different features of language complexity (e.g. de Villers and Pyers 2002);
3) Might be better identifiable in the archaeological record, as certain social behaviours are shown experimentally to require theory of mind (e.g. Remmel and Peters 2009);
4) The use of theory of mind does not limit interpretations to rare and recent timeframes, but can be applied throughout analyses of past social behaviour (Cole 2015)

This thesis uses a chain of inference from social learning, to theory of mind, to language, to construct a methodology for interpreting cognition and language from the Palaeolithic record. An inferential chain is only as strong as its weakest link, but it makes this argument fallable; this is regarded as a strength rather than a weakness, as theories should be scientific and able to be disproven or changed in consideration of new evidence. Wynn and Coolidge (2009, p. 119) propose a strict standard to make a rigorous argument in cognitive archaeology, which this thesis adheres to:

1) Cognitive validity: The evidence must actually require the abilities attributed to it. The cognitive ability must be one recognized or defined by cognitive science; it must be required for the actions cited; and the archaeological traces must require those actions. A strict standard of parsimony must apply. If the archaeological traces could have been generated by simpler actions, or simpler cognition, then the simpler explanation must be favoured;
2) Archaeological validity: The archaeological evidence must itself be credible. The traces in question must be reliably identified and placed appropriately in time and space.

In addition to this, Wynn (2009) has also noted the importance of the inclusion of psychological knowledge for developing robust hypotheses about cognition from archaeological material, and efforts have been made here to ground the theory firmly in the psychological literature.

This type of approach, which in this thesis marries theory to actual archaeological applications with interpretive, falsifiable results, is necessary for progress in the discipline of evolutionary cognitive archaeology (that is, the sub-discipline of archaeology concerned with the evolution of human cognition). As yet, the discipline lacks an empirically supported method for assessing either cognitive or linguistic ability in the Palaeolithic (Wynn and Coolidge 2009). What exists instead is aforementioned focus on theoretical discussions of Palaeolithic 'symbolic' objects (e.g., Marshack 1976, Chase and Dibble 1987, Lindly and Clark 1990, Davidson 1991, d'Errico et al. 2003, Bednarik 2008, Henshilwood and Dubreuil 2009, Vanhaeren et al. 2013, Burdukiewicz 2014, Langley 2015).

1.5 Justification of Method

This thesis will form its inferential chain through a number of areas of scientific literature, making it an interdisciplinary work. At the forefront of the link between language and theory of mind is the child developmental literature, through which researchers have understood the ontogeny of these intertwined abilities. The medical and educational literature is also of importance, as the contribution of theory of mind training and autism studies are also of value to understanding the relationship between language and theory of mind. Comparative biology of non-human species is also important to understanding the evolutionary uniqueness (or not) of certain cognitive and linguistic abilities.

The link between theory of mind and social learning will therefore be built using literature from the areas of child development, cognitive science, and comparative biology, which has a large focus on how different species transmit and acquire certain cultural features through social learning. Whereas the link between language and theory of mind is more established, the link between social learning and theory of mind lacks the levels of empirical investigation especially in higher orders of theory of mind, and relies more in theoretical and parsimonious linkages.

As this thesis is investigating basic levels of theory of mind, ethnographic examples of teaching and learning, such as recorded in studies of apprenticeship in modern human groups (Stout 2002), will not be explored. Modern humans are capable of complex theory of mind and culture such that extraneous variables involved in the creation and replication of material culture will be too difficult to tease out, and this is a prime reason for the choice of an experiment that simulates and controls for certain cognitive features and replication environments.

1.6 The Experiment

This thesis has been motivated by the need for an empirical method of assessment to infer hominin cognitive abilities from Palaeolithic material culture. Because of the often noted 'ephemeral' nature of language (see quotations listed at the beginning of this chapter), an experimental approach is ideal for generating data concerning the origin of language.

An experiment was conducted to test whether different social learning methods create differing morphological characteristics in stone tool assemblages. Twenty novice knappers in four different simulated social learning environments attempted to replicate a model handaxe a number of times after learning in their group-specific way, which simulated a specific social learning method (in this case emulation, imitation and teaching) which required different levels of theory of mind (simulating no theory of mind, second-level intentionality, and third-level intentionality). As these groups differ in their use of theory of mind to acquire knowledge about replicating the handaxes, the impact of differing levels of theory of mind is explored in a morphological analysis of the resulting four assemblages of tools and their debitage, to explore the effects of their different conditions. As stated in the hypothesis, it is predicted that these simulated social learning groups will produce assemblages with differing ranges of morphological variability according to different measurements, and this is hoped to be used to identify complex social learning methods in the archaeological record.

1.7 Chapter Outlines

This chapter has thus far presented the research question, the hypothesis, the rationale for this thesis, its theoretical viewpoint and methodological justifications, and the experiment. It has briefly outlined why different social learning methods would indicate the theory of mind ability of the tool-maker, and why indications of theory of mind ability would therefore indicate linguistic ability, creating a chain of inference (Figure 1.1). The following chapters will examine topics in greater detail.

Chapter 2 will explore the interdependent relationship between language and theory of mind abilities. It begins with a definition of theory of mind and levels of intentionality, includes a brief history of its study, and then describes its developmental relationship with language in normally developing children, adults, and people with impairments (both cognitive and physical). The chapter will also investigate theory of mind and language in other social animals such as non-human primates and avian species. The chapter's aim is to establish the predictive relationship between language and theory of mind such that this relationship can be extended to pre-modern hominins, and show how a co-evolution of the two abilities is likely. A model of the co-evolution of theory of mind along with what linguistic features would present itself will be illustrated, including word reference correlating with second level intentionality, and mental state terms and complementation with third level intentionality (levels of intentionality are explained in Section 2.2). A co-evolutionary narrative is then presented of the development of early theory of mind and language together.

Chapter 3 will explore social learning, and the role theory of mind has to play in how information is transmitted culturally through different social learning methods such as stimulus enhancement, emulation, imitation and teaching, in our own and other species. It will describe the research that shows the role of cultural transmission's effect on material culture in terms of copy error, alluding to the construction of this thesis' experiment. It will also examine recent work on cultural transmission and its contribution to our understanding of how material culture spreads and changes due to how people teach and learn. It will establish how theory of mind is implicated in different social learning mechanisms such that if researchers find evidence for specific social learning mechanisms in the archaeological record, that would be indicative of theory of mind as well. It then critically analyses the few other knapping experiments that have examined social learning, but which have not extended the results to appreciate their full cognitive implications (although in some cases the author argues they have taken implications for silent teaching as a proxy for language-less hominins too far).

Chapter 4 will present the design of the experiment and the methodology used. It details the ethical considerations, a detailed experimental framework including information on the participants, the groups, and the experimental materials. Expectations of the experiment are also outlined, before detailing the data collection procedure, the recording methodology, and the analytical procedure for the experiment, so that the reader has a full understanding of the motivations behind its setup and the choices made therein.

Chapter 5 will then present the results of the experiment. It begins by outlining which data was included in the analysis, before presenting the results from the traditional morphometric variables involved to show knapping intensity and reduction (weight lost, scar count and scar density), and flake attributes (flake count and flake weight). A geometric morphometric analysis of handaxe shape follows, where ordinations are presented along with significance testing of the groups to show where they can be successfully statistically discriminated.

Chapter 6 presents the discussion of this thesis' results in light of the research question. It will contextualise the support it offers alongside results in related studies that also look at social learning through knapping experiments. The interpretive potential, as well as limitations of the thesis will also be discussed. Key to this chapter will be a discussion of the integration of the research framework with the archaeological evidence, where broad technological industries (Lomekwian, Oldowan, Acheulean, and Levallois) will be interpreted alongside the theory of mind and linguistic abilities the thesis predicts they present evidence for.

Chapter 7 will conclude this thesis. It will assess its contribution to the field, and discuss future work and potential experiments in situating cognitive and linguistic abilities in the material record through experimental knapping methodologies (which benefit from using porcelain as a controlled knapping material). Its final remarks reflect on the research question and the experimental evidence which supports that social learning is indeed a variable that affects the morphological variability of stone tool assemblages, and therefore assessing that variability can lend itself to interpretations in the archaeological record not only of the social learning method used by its makers and propagators, but the theory of mind necessary for the materials to persist in the archaeological record, and the level of linguistic ability needed to support that cognitive complexity.

1.8 Conclusion

This thesis makes the case that it is not language that is the condition for the transmission of complex stone tool technologies, but *theory of mind,* of which language is a condition for its development; it argues that higher levels of theory of mind are unattainable without the developmental support provided by language. In this way, language *can* be seen as a condition for the transmission of lithic technology, when technological transmission is of such a complexity that theory of mind is necessary for its successful transfer. In order to interpret language from Palaeolithic material, this thesis tests whether lithic technology learned by emulation, imitation and teaching is morphometrically distinguishable. Identifiable traits at the assemblage level of morphological variability can contribute to the identification of different methods of social learning used by Palaeolithic hominins through the cognitive abilities they require.

Following the results of the experiment, the argument will be that low morphological variability within lithic assemblages (high standardisation) signal high fidelity social learning, theory of mind, and language abilities. By offering theory of mind as a proxy for language ability in the Palaeolithic record, with the support of an empirical experiment, this is the thesis' original contribution to knowledge in the discipline. This conclusion crucially lends itself to further empirical research in the archaeology of language origins. Including an analysis of archaeological material as a case study was deemed beyond the scope of the current project (considering time and length constraints), but this thesis permits the future investigation of the archaeological material, and must be preceded by the research conducted herein before the necessary archaeological applications can be made.

The experimental approach used in this thesis, which generates vital data about the learning of stone tool technologies, is essential to the progress of cognitive archaeology, evolutionary linguistics, and to anthropology as a whole, as they lack empirical approaches to questions concerning the evolution of language. Archaeology is the only direct access to the context and material associated with the people for whom language and human-like cognition developed. If it is in fact found that the presence of language can indeed be identified through stone tools, language can, in a sense, be said to fossilise.

Chapter Two.

Theory of Mind as a Proxy for Language

2.1 Introduction

The main argument of this thesis relies on adequately presenting through previous studies how theory of mind and language are intimately related cognitive abilities, and have a predictive value which can approximate the ability of the other; theory of mind only develops with the scaffolding provided by certain linguistic structures, and likewise language only develops with the scaffolding provided by certain theory of mind and pre-theory of mind abilities (such as gaze following and joint attention). This thesis will then argue that the predictive quality of language and theory of mind allows for assessing the ability of the other in an evolutionary scenario. While some researchers have already argued for a co-evolution of language and theory of mind (e.g. Malle 2002), this thesis is unique in suggesting that theory of mind can be identified archaeologically in order to support the existence of language ability. The research question presented in Chapter 1 is built upon a chain of inference where social learning necessitates theory of mind, which implies language ability. This chapter will build a case for the link between language and theory of mind, such that they have a predictive relationship which allows theory of mind to act as a proxy for language ability in the archaeological record.

This chapter will first examine the cognitive ability called theory of mind and its definition, including levels of intentionality. It will then review a history of its study and the consensus view of its development in children. It will discuss theories of its representation and of its neural processing, before introducing its correlation with language ability. A review of a number of important studies follow, which present theory of mind and language's relationship in individuals with autism and for individuals that develop with hearing impairments. Studies of adult correlations and higher order theory of mind are discussed, as well as the impact of language training. Studies involving non-humans are also discussed, where the focus is mainly on chimpanzee studies (as they are the species most examined in this aspect). These studies contribute to the argument presented, and are echoed by other researchers, that theory of mind and language share a co-evolutionary origin. This culminates in a narrative of the co-evolutionary process of language and theory of mind, including a model presented in Figure 2.2. At the end of this chapter, sufficient evidence will have been put forth that allows for the link between theory of mind and language ability's predictive relationship to be established. It follows then that theory of mind can be used as a valid proxy for Palaeolithic language ability, which could prove easier to identify in the archaeological record than direct signs for language itself.

2.2 Definitions

Theory of mind is the ability for individuals to theorise about the mental states of others (Wellman et al. 2001),such as their desires, intentions, knowledge and beliefs, and predict how they motivate the behaviour of others (whether those interpretations being made are correct or not). As others' minds are not directly accessible or observable from just bodily cues, individuals must infer the mental states of another's mind by theorising about these processes, which is why the term is referred to as 'theory' of mind. With this ability, an individual can attribute others, and themselves, as a thinking agent, and understand that individuals hold their own mental states that are independent of their own. Language, on the other hand, refers in this thesis to a system of symbols used for intentional communication. While definitions for what language is and is not abound, such as Hockett's design features (Hockett 1960, and for a more in depth discussion see Aitchison 2008), this thesis uses a broad definition that differentiates language from other species' communication systems (Burling 1993) by highlighting those features that make it a qualitatively human ability; these are judged here to be its symbolism, and its intentionality. As it is a broad definition that does not delve into the intricacies and finer points of grammar or innateness, it is expected to be an uncontroversial definition for this thesis.

Theory of mind (often abbreviated as ToM) has been referred to synonymously as 'mentalizing' (Frith and Frith 2003), 'mind-reading' (Dawkins and Krebs 1978), and 'folk psychology' (Olson and Bruner 1996). It is equivalent to what is termed second-order intentionality (Dunbar 2004). Intentionality is a philosophical concept defined by Daniel Dennett (1983) as 'aboutness', and where levels of intentionality grade higher when one mental state is embedded within another. This process of a constituent embedded in a like constituent (called recursion) means theory of mind can embed thoughts about thoughts about thoughts (Figure 2.1). First-level intentionality could be a way to describe what could be broadly term consciousness, with the ability to 'intend', but not necessarily consider own or others' intentions (whereas zero-level could be imagined to be pure 'experience'). Second-level intentionality is therefore considered theory of mind, as this is the level at which thoughts are about thoughts (the point which introduces a hierarchy). This embedding enables an individual to represent a false belief, i.e. understand that another may perceive a situation in a different way, which may or may not reflect one's own interpretation of reality (Doherty 2009). Language, which will be discussed here in its relation to theory of mind, is also recursive in that its hierarchical organisation system (syntax), is what gives individual elements (morphemes, or what can roughly be considered words) greater meaning with their constructed relation to each other. Having a theory of mind is significant socially, because it allows for a sharing of mental worlds. Once an individual attributes others with like minds, a new mental environment is accessed where thoughts, feelings and knowledge can be shared and appreciated by others. New emotions which depend on thinking about the thoughts of others, such as embarrassment or guilt, suddenly become possible and take the potential social relationships and interactions to a new level. As Miller (2009, p. 749) writes, those with theory of mind can now "manipulate others' mental states in order to tease or deceive; more positively, they can console, cooperate, and in general coordinate their actions with the beliefs and desires of their interaction partners." The attainment of theory of mind is the point at which mental landscapes are opened and populated.

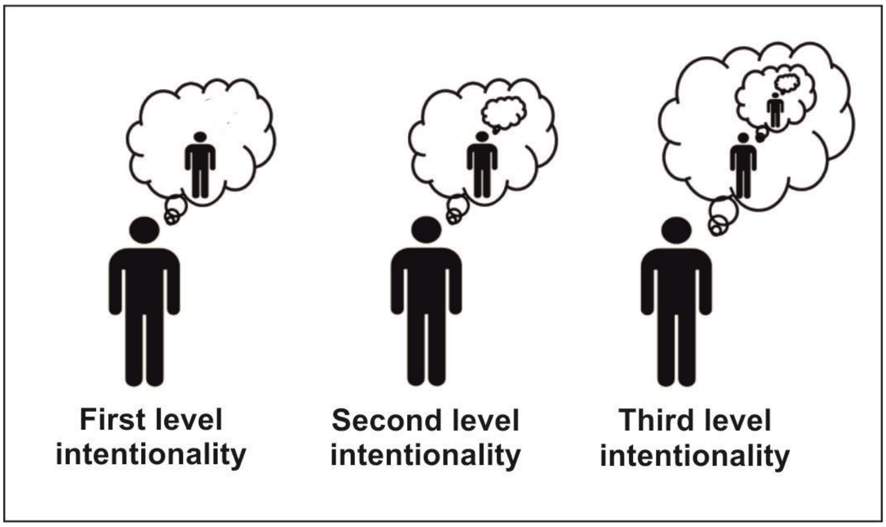


Figure 2.1 Levels of intentionality, where thoughts about another's thoughts can be embedded in a hierarchy of increasing complexity.

With thoughts containing third-level intentionality (also confusingly called second-order beliefs, Astington et al 2008), researchers are able to confidently speak about cognition that is exclusive (in the present day at least) to modern humans (considering the difficulty in establishing that non-human species are able to process thoughts using second-level intentionality, Call and Tomasello 2008). At third-level intentionality, an individual *thinks about what another thinks about what another thinks* (e.g. Baron-Cohen et al, 1997). It opens up the mental landscape further, in that it allows relationships to go beyond dyadic two-person sharing of mental worlds, and to appreciate external relations. This is because in a thought process where an individual can think about what another thinks about themselves, social pressure can be present (such as complex forms of bullying, Sutton et al. 1999). On this mental landscape, social rules can be formed and enforced.

In an example of humans' use of fourth-level intentionality, Dunbar (1998) notes that composing a novel requires at least this level of intentionality, because the writer must predict what the reader understands a character thinks about what another character thinks. As Wellman et al. note (2001, p. 655), "Because actors have certain desires and relevant beliefs, they engage in intentional acts, the success and failure of which result in various emotional reactions." An author of a piece must then always be able to comprehend a higher level of intentionality than the reader is expected to, in order to construct the elements of a story and manipulate the expectations of the audience.

Levels of intentionality over fourth- or fifth-level are shown to be difficult for humans to engage with, and it seems to not simply be a factor of memory constraints; cause-and-effect chains concerning objects have been shown experimentally to not be under the same cognitive limitations (Dunbar 1998). This difficulty of holding in mind very complex theory of mind operations is illustrated with an example from Dennett (1983, p. 345):

"...I suspect that you wonder whether I realise how hard it is for you to be sure that you understand whether I mean to be saying that you can recognize that I can believe you to want me to explain that most of us can keep track of only about five or six orders [of intentionality], under the best of circumstances."

For the scope of this thesis, the focus will remain on levels of intentionality 1-3, which are levels associated with the relevant social learning mechanisms for the experiment outlined in Chapter 4 (emulation, imitation and teaching).

2.3 History of Study

Theory of mind was first used as a term in a classic paper, "Does the chimpanzee have a theory of mind?" (Premack and Woodruff 1978). In this study, the chimpanzee Sarah was subject to a number of tests to see if she understood human goals by being shown videos of a human struggling with simple problems. The chimpanzee was then offered photographs of the scene, one with a possible solution to the problem. The authors concluded, "The chimpanzee's consistent choice of the correct photographs can be understood by assuming that the animal recognised the videotape as representing a problem, understood the actor's purpose, and chose alternatives compatible with that purpose." (Premack and Woodruff 1978, p. 515). However, whether or not chimpanzees, or any other non-human species, are capable of theorizing about the mental states of others is inconclusive. Research seems to have shown apes capable, in some instances, of being able to consider the goals and intentions of others. However, they have not been successful in non-linguistic false belief tasks (Call and Tomasello 2008, Penn and Povinelli 2007, but see Krupenye et al. 2016 and Buttelmann et al. 2017). Whether or not that is because they do not have theory of mind, do not have another necessary cognitive feature that causes them to fail the tasks, or it is due to a shortcoming of the task structure (the difficulty in setting up non-linguistic false belief tasks has been noted by researchers) is unclear. I further discuss the potential for theory of mind abilities in enculturated chimpanzees and other non-human species in Section 2.13.

Following Premack and Woodruff's (1978) paper, theory of mind ability was investigated in the cognitive development of human children (Bretherton et al. 1981; Wimmer and Perner 1983). Wimmer and Perner (1983) developed the 'false belief task', which has become the standard way of assessing theory of mind ability in experimental settings. There are many iterations of the false belief task (e.g. Call and Tomasello 1999), but they all aim to evaluate if a participant can appreciate that another agent can appreciate another has a belief or knowledge state that is different from their own. For example, in the Sally-Anne test (e.g. Baron-Cohen et al. 1985), a participant is shown (either through actors, puppets or pictures) two characters (who are often named Sally and Anne). Sally hides an object in location A, out of sight, and leaves the room. Anne then takes the object out of its hiding place and hides it in location B. The participant is then asked, "When Sally returns, where will she look for the object?", if the participant appreciates that another can hold knowledge that differs from their own perception of reality, they will answer that Sally will look for the object where she left it, in place A. However, those who struggle to hold a false belief might answer that Sally will first look where Anne moved the object to, as they do not see Sally as having a different perspective and knowledge state to themselves. Another example of tests for assessing theory of mind can be found in the appendix of Happé (1994). Sample stories and questions for higher levels of theory of mind ability are listed in Liddle and Nettle (2006). The issues associated with false belief testing (such as tests being failed due to other factors, and the problems of designing non-lingusitic tests) are dealt with in Leudar and Costall (2009), but it remains the most popular method of assessing theory of mind ability (Wellman et al 2001).

Theory of mind studies have roots in a number of cognitive impairments, and its understanding grew out of explorations of these developmental deficits and medical complications. Individuals diagnosed with Autism Spectrum Disorder, which affects social interaction, are found to have symptoms associated with deficits in the development of their theory of mind ability, a theory put forth by Simon Baron-Cohen (Baron-Cohen et al. 1985; Baron-Cohen 1995; Baron-Cohen et al. 1999; Baron-Cohen and Wheelwright 2004). Theory of mind is also associated with a number of other conditions including schizophrenia and psychopathy (Brüne and Brüne-Cohrs 2006), depression (Bora et al. 2005), attention deficit hyperactivity disorder (Korkmaz 2011), and a long list of other behavioural and clinical disorders (Sprung 2010). These studies are vital to appreciating the contribution theory of mind has to social behaviour, and the effect that an impaired theory of mind has to the development of other abilities, such as language. A number of them will be discussed in this context in Section 2.6.

2.4 Child Development

The typical path of development for children as they acquire language and theory of mind has not yet been outlined. As part of this thesis, I argue that language and theory of mind not only correlate in ability in modern populations, but it is likely that these abilities co-evolved and therefore correlated in Palaeolithic populations as well, including ancestral species. I have waited until now to describe the human ontogeny of these two abilities; language and theory of mind share a close development and have similar pre-adaptations (skills that scaffold or allow the acquisition of another), which makes it logical to discuss them at the same time. It is the scaffolding and influence that I suggest makes language and theory of mind to have likely evolved together, resulting in their predictive qualities to be used for this thesis' purpose of promoting theory of mind as a proxy for language in the Palaeolithic. Here I will describe the development of theory of mind and language in typically developing human populations, followed by specific examples of where one scaffolds the acquisition of the other. This will be in preparation of the following section, where I will formally lay out a model of the co-evolution of theory of mind and language in a series of interwoven stages where each feature scaffolds the development of the next.

Language and theory of mind development are preceded by the development of a number of abilities during infancy. Infants must learn the difference between animate and inanimate objects (Gelman and Spelke 1981), and learn to understand facial expressions, and reference their parents' faces for cues on how to approach a novel object. They learn to follow the gaze of another, associating it with points of interest (Brooks and Meltzoff 2008). They learn that a person has a different line of vision from them, and how that line of vision might differ in what they can see, which is called visual perspective taking (Hamilton et al. 2009). These skills lead to an ability called joint attention, where an individual appreciates shared attention of an object (Bruinisma et al. 2004, Toth et al. 2006). Joint attention allows gesturing of intention such as pointing (Gómez 2007). Figure 2.2 shows a number of aspects of theory of mind and the age at which they are acquired in typically developing children (from Miller 2006), but many of the emerging abilities allow for the acquisition of language abilities as well.

Infants typically learn first words from about ten-to-twelve months (Baldwin and Moses 1994), around the time they are beginning to be able to tell where others' attention is fixed. Bruinisma et al. (2004, p. 993) call joint attention a "cluster of behaviors that share the common goal of communicating with another person about a third entity in a non-verbal way, including eye gaze alternation and gesturing." Charman et al. (2000) found that joint attention positively correlated with later theory of mind performance from two-to-four years old. Joint attention emerges in normally developing infants at around nine months (Miller 2006). It is a mentalising ability that brings the individual closer to theorising about the mental states of others, and without it, the linguistic system of tying a word to a referent would never come to be shared between two individuals. This is an important way that theory of mind ability is fundamental and necessary for the acquisition of certain linguistic structures, as a lexicon could never emerge without joint attention and the ability to appreciate a common symbolic reference for an entity.

Age	Aspects of theory of mind
6–12 months	• Joint attention, including gaze and point following, and alternation of gaze between person and object (Bruinsma et al., 2004; Carpenter et al., 1998) • First words (Tomasello, 1995)
13–24 months	• Recognize intentionality in others as demonstrated in word use (Tomasello, 1995) • Recognize that others have desires different from one's own (Repacholi & Gopnik, 1997) • Early pretend play (Leslie, 1987)
30–36 months	• Begin to use mental state terms with truly mentalistic functions (Bartsch & Wellman, 1995) • Increasingly sophisticated pretend play (Youngblade & Dunn, 1995)
37–48 months	• Increasing ability to understand how things look from another's perspective (Flavell et al., 1981) • Begin to understand sentence complements (de Villiers & Pyers, 2002)
49–60 months	• Consistently pass false belief and appearance-reality tasks (Wellman et al., 2001)

Figure 2.2 A developmental timeline of key stages of theory of mind and language acquisition (from Miller 2006).

Around the time of a child's first birthday, they are beginning to understand other people as not only intentional agents, but mental agents too (Tomasello 1995). In Träuble et al. (2010), experimenters used a non-verbal violation-of-expectation task, where they considered fifteen-month-old infants' looking times when an outcome was unexpected. When an adult actor acted out failing the test, by reaching for the box without a reward, infants looked longer than if the actor selected the 'correct' box, as if they knew that the adult's goal had not been met.

Between thirty-to-thirty-six months, children begin to use mental state terms, such as want, think, and know (Bartsch and Wellman 1995). But it is between the three-to-five year old period that children go through a change where they comprehend that others may hold beliefs that differ from their own, or a false belief. Wellman et al. (2001) present a meta-analysis of false belief testing on young children, and conclude that "understanding of belief, and, relatedly, understanding of mind, exhibit genuine conceptual change in the preschool years" (Wellman et al. 2001, p. 655). While a number of studies point to infants having theory of mind-like abilities before this threshold time, it is accepted in the field that theory of mind is acquired in children between ages three and five.

2.5 Representation and Processing

Much discussion of theory of mind has centred on how it is processed (Leslie et al 2004, Baron-Cohen and Cross 1992). Two main accounts of theory of mind representation have been proposed (Carruthers and Smith, 1996), and both regard cognition as very internalist, mentalist operations. 'Theory theory' (Stich and Nichols 1992) posits that people apply theory of mind in a theory driven way, where predictions about mental states are used to explain others behaviour, like a set of developed laws. Some researchers have posited a specific theory of mind 'module' for the processing of this ability (Scholl and Leslie 1999), or explained it as a number of domain-general skills working together (Gerrans 2002). 'Simulation theory' on the other hand (Gordon 1986), questions this representation, and suggests that people analogise or 'simulate' their own processes in order to predict the mental states of others, in a sort of mental imitation. The positive point of simulation theory is that it provides a mechanism where predictions are made about others' mental states by learning about one's own, and provides a less cognitivist account of theory of mind representation, which could be more compatible when taking into account the role of social interactions and material engagement, which also influence and constitute an important part of the cognitive process, a role left out of most discussions of theory of mind representations which consider theory of mind as 'all in the head'.

Much research has also been conducted on the neurological regions involved in processing theory of mind. This has included studies that look at the areas involved in reflecting on one's own mental states, the mental states of others, and in the processing of both (Abu-Akel 2003). The limbic-paralimbic system such as the amygdala, as well as the prefrontal cortex, have been recorded as involved in processing both one's own, as well as others' mental states. In addition, when considering the mental states of others the superior temporal sulcus is also activated. When people consider their own mental states, the inferior parietal lobe (which hosts abnormalities in those with autism and schizophrenia), part of the right posterior parietal region, takes a specialised role (Vogeley et al. 2001).

Regardless of the representational processes or neurological regions governing its processing, theory of mind describes a set of cognitive behaviours of which great import must be placed due to their role in such key human faculties. Part of the problem of their explanation is similar to the mechanistic description of many forms of representation. The theories of representation and processing of 'theory theory' and 'simulation theory' speak of theory of mind 'in the brain', and this might be a barrier to its full understanding, since theory of mind operates at a social level in the environment, engaging with people and materials. The position taken in this thesis is that the mind is both extended and distributed, and this is important for how ancient hominins engaged with materials (Malafouris 2013). It is not a reason to dismiss theory of mind as incompatible with other theoretical frameworks, but to be optimistic that future studies will elucidate the workings of the mind, perhaps by incorporating what is known about its interactions with environments and materials further. The semantics behind the labels describing its representation can differ, but the behaviour being described is the same, and all are compatible with the hypothesis of this thesis.

2.6 The Correlation with Language Ability

This notion of a predictive relationship between language and theory of mind comes primarily from child development research, where educational and medical researchers have long attributed importance of the role of theory of mind in language learning and *vice versa* (Miller 2006, Garfield et al. 2001, Jenkins and Astington 1996). Much of this correlation was first discovered through autism research, where in autism spectrum disorders (ASD) a key trait is impaired or delayed theory of mind (Baron-Cohen et al. 1985, Baron-Cohen 1995). Language and theory of mind are deeply interconnected by their very nature (Astington and Baird 2005; de Villiers 2007). Language is a system of culturally learned symbols (Aitchison 2008) which are often used to intentionally convey mental states (knowledge, belief, emotions, desires, intentions), which is fundamentally grounded in theory of mind, as "we cannot make sense of communicative interactions without presupposing that the interlocutors possess mutual knowledge of relevant beliefs and intentions." (Baldwin and Moses 1994). Language users also need to appreciate the behaviour of an agent as intentional to perceive a specific behaviour as communicative. Theory of mind is the ability to attribute mental states to another in order to understand social interaction, necessary for the perception of behaviour as being driven by a mindful agent. Theory of mind enables the explanation and prediction of behaviour by "theorising" about the content of others' beliefs, desires or emotions. Like language, some researchers argue that it arose to cope with an evolutionary scenario where social skill contributed to fitness (Brüne and Brüne-Cohrs, 2008; Dunbar, 2003; Byrne and Whiten, 1988). It is this close relationship and overlap in social function between the two abilities that forms the theoretical foundation of this project, promoting theory of mind as a proxy for language ability in the archaeological record.

Language and theory of mind both develop in the early years of life, and their function and further development are fundamentally interdependent. The most interesting aspect that pertains to language evolution studies and the argument constricted in this thesis is language and theory of mind's strongly correlative relationship, where an individual's ability in one predicts their ability in the other according to standardised test scores in adults, developing children and infants, and individuals with cognitive or physical impairments (Astington and Jenkins 1999, Courtin 2000, Cutting and Dunn 1999, Dahlgren et al. 2010, de Villiers 2005, Ebert 2015, Eisenmajer and Prior 1991, Farrar and Maag 2002, Farrar et al. 2009, Happé 1995, Hughes 1998, Hughes and Dunn 1997, Jenkins and Astington 1996, Milligan et al. 2007, Rakhlin et al. 2011, Ruffman et al. 2003, Schick et al. 2007, Slade and Ruffman 2005; Tager-Flusberg and Joseph 2005). The predictive relationship of these two abilities also exists universally in modern populations (Avis and Harris

1991, Chasiotis et al. 2006). Language has been suggested to be “a necessary ontogenetic precursor of mature theory of mind.” (Segal 1998, p. 155), whereas without theory of mind, there is no grounds on which intentional communication can work in the first place. Language and theory of mind’s reliance on each other during development is the reason which some researchers posit a co-evolution of the two abilities in our species (Corballis 2012, Malle 2002, Worden 1998, Smith 1996, Shatz 1994); Malle (2002) speaks of language and theory of mind as inseparable, and supports an “escalation process in which language and theory of mind have fuelled each other's evolution.” (p. 265).

I will argue here that different levels of theory of mind imply the possession of specific semantic (what linguistic elements mean) and syntactic (how linguistic elements are organised) features that are necessary for theory of mind’s ontogenetic development. The two abilities develop in tandem in children, one fundamental to an increase in complexity of the other (Miller 2006, de Villiers 2007, Garfield et al. 2001). Second-level intentionality (basic theory of mind) develops only with the possession of basic linguistic skills, such as understanding and using words (Miller 2006). This requires being able to conceptually anchor a learned symbol to an object, and appreciate its shared meaning; in effect, “naming” something. Understanding when someone is focussed on an object, and that the object is associated with a word, aids in the understanding that others possess minds, goals and intentions, and that those mental states can be shared and communicated. Theory of mind supports the acquisition of language in this way, but language also supports the acquisition of more complex levels of intentionality: Third-level intentionality implies the possession of more complex linguistic abilities, including the semantic labels and linguistic architecture to attribute mental states to others linguistically. These are skills that are developing during the acquisition of second-level intentionality, but are integral to the acquisition of third-level intentionality (Lockl and Schneider 2007). The linguistic features to attribute mental states are:

1. Mental state verbs (want, think, know, believe etc.);
2. Complementation (where a linguistic constituent requires another to complete its meaning, Crystal 1995).

Complementation is a simple grammatical relation made up of a head and a complement (Figure 2.3), of the sort that is necessary to attribute an agent with a quality, such as a person and their desire or a person and their belief. These are both shown to enhance theory of mind appreciation (de Villiers and Pyers 2002). These linguistic skills are necessary for an individual to access not only the minds of others, but to understand that others' minds theorise about yet others' minds as well.

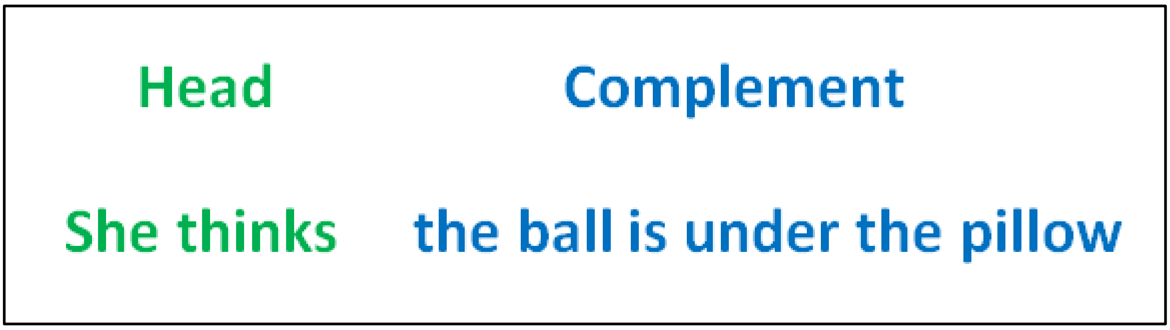


Figure 2.3 An example of complement structure, where the head requires the complement to complete its meaning.

The relationship and developmental interplay between language and theory of mind is not surprising, considering their functional overlap. Pragmatics, a sub-discipline of linguistics, is essentially the attribution of theory of mind through linguistic means, how for example the utterance "I'm cold," can be said with the purpose of encouraging someone close a window.

2.6.1 Directionality of Influence, and the Contributions of Syntax and Semantics

In many of the early studies comparing theory of mind and linguistic ability, a key research question was investigating the directionality of developmental influence: was theory of mind supporting language development, or was language supporting theory of mind development? Astington and Jenkins (1999) published a study to explore the directional contribution of theory of mind to language ability and vice versa, as well as examine the different contributions of syntax and semantics. A longitudinal study tested fifty-nine three year olds over seven months, testing them at three different points over time. They found that language ability predicted their later theory of mind performance, but that their theory of mind ability did not predict their later language performance. They also found that the prediction of language scores to later theory of mind scores were most predicted by syntactic abilities, rather than semantic ones. They noted that theory of mind scores not predicting later language scores might have partly been due to the range and variance of the language test scores being greater than the theory of mind test scores. By collapsing the scores from the standard language test into quartiles, the range and variance was reduced. Still, they found a similar result, though were still wary due to other psychometric issues. They therefore concluded that their research is consistent with studies that show that "language plays a fundamental role in theory-of-mind development." (p. 1319) While directionality had not been confirmed, it was uncontroversial the abilities strongly correlated.

In a study in 2005, Slade and Ruffman tested forty-four three-to-four year old children who were given three false belief tests, a working memory test, and four language tests for syntax and semantics skill, and then six months later were tested again. They were cautious about the findings in Astington and Jenkins (1999) that language ability supported later theory of mind but not the other way around, because of other evidence they were aware of that precursors to theory of mind contribute to language understanding. They also wanted to test Astington and Jenkins' finding that syntax contributed to later false belief understanding more than semantics did, and they did so by using two tests that specifically focussed on syntactic skill, and two that specifically focussed on semantic skill. In their conclusions, Slade and Ruffman found a bi-directional relationship between theory of mind and language, but only after adjusting the sampling and using a false belief composite score, solving some of the psychometric issues Astington and Jenkins had previously alluded to in their studies. Slade and Ruffman did not find that syntax played any more of a role than semantics in language's contribution to theory of mind when separated into different tests and with the composite score from false belief ability; both were found to have a predictive value for later theory of mind. They also found that working memory (immediate conscious perceptual processing) was not a predictor of later false belief ability or *vice versa*.

Farrar et al. (2009) wanted to assess the relative contributions of a number of linguistic attributes on understanding theory of mind. They tested thirty-four preschool children with specific language impairment (SLI) and found that grammatical and vocabulary development both contributed to theory of mind reasoning. However, their findings disagreed with that of de Villiers and Pyers (2002) and others that attributed a special role to complementation to the acquisition of theory of mind abilities, arguing instead that it was general grammatical ability as a whole, at least in children with SLI. But in yet another study looking at the relative contributions of different aspects of language to theory of mind ability, Rakhlin et al. (2011) found that those who succeeded on the false belief tasks had significantly higher syntactic complexity scores, in a study of fifty-four children aged between six-to-twelve who had a language disorder and varying language skills.

As these studies were interested in investigating the acquisition of basic theory of mind, or second-level intentionality, it would be interesting to see the relative role of different linguistic components on more advanced applications of mindreading and if that involved a differential contribution from specific linguistic components, as studies by de Villiers and Pyers (2002) implies. Studies of adult theory of mind and language, as well as higher order theory of mind and language will further elucidate the directionality of influence of aspects of language and theory of mind. The relative contributions of syntax and semantics have conflicting results in a number of studies, but it is clear that both make a contribution, and it could be that one makes a larger contribution than the other in certain situations rather than others.

2.6.2 Autism Studies

The link between autism and theory of mind was first examined by Baron-Cohen et al. (1985), which followed with the account that autism can be explained as a communicative and social impairment of which theory of mind is a key aspect. Different theory of mind studies with autistic children reported that the majority of autistic children did not pass false-belief tasks (Happé 1995). In one instance, Happé (1995) investigated the role of verbal ability in the success of autistic individuals in passing false belief tasks. She looked at data for seventy autistic children, thirty-four mentally handicapped (but non-autistic) children, and seventy normally developing children across different ages. All individuals had been given two false belief tasks, including memory and reality questions as controls. Children had been tested with the British Picture Vocabulary Scale, which gives a score of 'verbal IQ' and 'verbal mental age'. The proportion of participants that passed or failed the false belief tasks did not differ between the normal and handicapped groups, but the autistic group failed significantly more often. For the autistic group, verbal ability was the only significant predictor of success on false belief tasks, while both age and verbal ability were significant predictors for the normal group. This makes sense, as normally developing children progress in their development at a predictable rate, but autistic individuals may be impaired to very different levels. Interestingly, the autistic group seemed to need much higher verbal ability to pass theory of mind tasks, which is suggested to possibly be due to the tests posing different problems to autistic and normally developing children. Another possibility given was that, "the high VMA [verbal mental ability] found in autistic subjects who pass theory of mind tasks is to suggest that these subjects are solving the tasks in a verbally mediated fashion. A number of authors have mentioned their impression that autistic subjects solve theory of mind tasks in an unusually conscious and logical way, for example, looking as if they are doing "mental arithmetic" before eventually giving the correct answer to second-order theory of mind tests." (Happé 1995, p. 852).

In another study by Bowler and Strom (1998), the researchers gave a number of children false belief tasks which the children had previously failed, but repeated with enhanced behavioural and emotional cues. The typically developing children, as well as the autistic children, both exhibited improved scores, while the control group, who had other learning disabilities, did not improve with the emphasised cues.

Autistic individuals are also shown to pass false belief tests especially in cases where the individuals are high functioning or have Asperger's syndrome, as autism is a spectrum disorder. In Paynter and Peterson (2010), sixty-three children aged between five-to-ten years old (in which there were typically developing, high functioning autistic, and Asperger syndrome participants), were tested in a number of false belief tasks that related to both semantic and syntactic skills. They found that the Asperger Syndrome individuals and the typically developing individuals did not differ in their theory of mind understanding. It is also interesting to note that while Asperger syndrome is on the autism spectrum, significant language delay is not one if its diagnostic criteria (Bennet et al. 2008). The high-functioning autistic individuals in the study however still experienced a substantial impairment to success in the theory of mind tasks. They also noted that syntax was a more reliable predictor of theory of mind skill than semantic knowledge, which they hypothesised might be the case as Kjelgaard and Tager-Flusberg (2001) found that syntactic development was a stronger indicator of how severe autism symptoms were than semantic development was.

2.6.3 Hearing Impairment Studies

Studies of the development of children with hearing impairment provide an intriguing case of varying speed of language acquisition due to varying levels of language input, and are therefore interesting to examine in regards to their correlating theory of mind development. In a diverse study of deaf children, Schick et al. (2007), tested the language skills, non-verbal intelligence skills, and false belief ability of 176 deaf children aged from three-to-eight years old who used either American Sign Language or English with either hearing or deaf parents. They wanted to know if language affects children's theory of mind ability because the tasks demand it, because of how it affects other cognitive processes, or because it is responsible for language acquisition. They found that deaf children of hearing parents had a significant delay in theory of mind acquisition, whether or not the parents knew American Sign Language. Deaf children of hearing parents typically experience language delay, as they are normally not brought up in as rich a linguistic environment due to less active and passive language input. However, deaf children of fluent signing households acquired language at the same rate as hearing children. Schick et al. (2007) found that vocabulary, as well as syntactic complements, were significant predictors of theory of mind task success, and that in all groups, language was a significant predictor of false belief score.

The findings of Schick et al. (2007) correlate with a previous study by Courtin (2000), who showed that sign language in deaf children promoted theory of mind, and that early access to sign language meant better acquisition. Deaf children born to deaf parents acquired theory of mind abilities at a normal rate, while deaf children born to hearing parents experienced both language and theory of mind delays.

In a study by Remmel and Peters (2008), thirty children with cochlear implants were tested alongside thirty children with normal hearing for language and theory of mind ability. They reported that the children with cochlear implants exhibited little to no delay in their theory of mind or language skills, but that their syntactic ability was more correlated to their time since the cochlear implantation, rather than their age at cochlear implantation. They concluded that cochlear implantation may benefit theory of mind because of the increased access to mental state language, which is commensurate with the hearing impairment studies mentioned above. These studies and others have led Hao et al. (2010) to conclude, "language impairment is probably the main cause of ToM deficits in deaf children of hearing families." (p. 1492)

2.6.4 Adult Studies

As most of the studies noted so far have had child participants (and most studies of theory of mind indeed use child participants), it is useful to look at results from some adult studies of the link between language and theory of mind to see how the correlation exhibits beyond childhood. The correlation between theory of mind and language ability does not just exist during initial development, but extends through into adult life. This means that delays in development are impactful, and in some cases do not regulate quickly.

Clegg et al. (2005) looked at adult outcomes of individuals with childhood developmental language disorders. seventeen men were assessed in their mid-thirties on their language, literacy, theory of mind, and memory abilities, along with their siblings who did not have a language disorder, to match for environment and shared family background. Their results showed that the adults with the language disorder had normal intelligence, but experienced worse social adaptation compared with their siblings and controls. They also experienced more schizotypal features. The authors concluded that childhood developmental language disorders contribute to significant deficits in theory of mind in adult life.

In a study of adults with hearing impairment, a 2010 study by Hao et al. showed that deaf groups of adults that lacked mental state language performed worse than the hearing group on tasks testing mental state understanding and advanced theory of mind. They also found that language ability was the only predictor for this mental state understanding, concluding that, "mental state language fosters ToM [...] Thus, they seem unlikely to be comprehended without a corresponding language specific to the mental domain." (Hao et al. 2010, p. 1499). These studies, though fewer in number than those focussed on children, show that theory of mind and language retain their intertwined relationship long after childhood or during their early development.

2.6.5 Higher-order Theory of Mind Studies

By far, most investigation in theory of mind considers second-level intentionality, or basic theory of mind typically acquired in preschool-aged children, and there have been fewer studies looking at the correlation between third-level intentionality and language (Miller 2009). The literature refers to more complex forms of theory of mind as 'higher-order', 'advanced', and 'second-order false belief' (which can be confusing, as second-order false belief is testing third-level theory of mind, and false belief tests second-order theory of mind). Research shows children typically acquire third-level intentionality between the ages of five-to-six years old. In the standard Sally-Anne false belief test, a modification can easily test for third-level intentionality if the researcher asks the participant where another agent *thinks* Anne *thinks* Sally placed the chocolate, for example (Perner and Wimmer 1985).

In the most relevant study of higher order theory of mind for this thesis, Lockl and Schneider (2007) conducted an experiment that looked at language, theory of mind, and metamemory in 183 children between the ages of three and five years old. The theory of mind tests included both first- and second-order. Their conclusion, from which the results showed a correlative relationship between language and theory of mind throughout, was that "we suggest that language is fundamental for children's ability to deal with both first-order and second-order belief tasks, but for somewhat different reasons. Language seems to be needed mostly because it supports the child to understand belief representation in first-order belief tasks and because a good deal of information processing is required in the testing situation itself in second-order belief tasks." (Lockl and Schneider 2007, p. 151)

2.6.6 Language Training and its Impact on Theory of Mind

The relationship between language and theory of mind is further emphasised in studies that show that during language training, theory of mind scores can be improved. Lohman and Tomasello (2003) conducted a study where 138 three-year-old children were exposed to one of three training conditions to see if their false belief abilities improved. Children who were exposed to sentential complement syntax improved, as did children who were exposed to discourse about deceptive objects. Children however did not improve when they were exposed to the third condition, which was deceptive objects without accompanying discourse. The results suggested to the researchers that language was playing a key role in the improvement of their theory of mind abilities.

Hale and Tager-Flusberg (2003) also explored the impact of language training on theory of mind in a study with sixty three-year-old children who had failed false belief tasks and sentential complement tasks. Groups were trained in either theory of mind tasks, sentential complements, or relative clauses (as a control). The group trained in sentential complements improved in both language and theory of mind post-tests; however, the group trained in theory of mind tasks only improved their false belief scores and not in the linguistic tasks. The relative clause trained group had no improvement in either theory of mind or sentential complements. The researchers concluded that sentential complements therefore contribute to the acquisition of theory of mind skills.

2.6.7 Non-human Studies

A number of studies have explored chimpanzees' ability to understand mental states (Premack and Woodruff 1978, Tomasello and Call 1997, Penn and Povinelli 2007, Call and Tomasello 2008), but not in terms of their correlated linguistic abilities, or attributing any amount of success to the amount of linguistic training that they had previously received. It would be fruitful to explore if theory of mind-like skills were shown to increase when apes and other species were trained linguistically, or had other cognitive training that led them to acquire explicit or implicit linguistic skills. At present, I am unaware of such studies. It is known that apes are sometimes able to deal in deceptive behaviour (Whiten and Byrne 1988), and are able to understand the different visual perceptions of others (Hare et al. 2000, but see Reaux et al. 1999), and might be capable of the imitation of certain behaviours (Whiten et al. 2009). In one study, Buttelmann et al. (2007) showed that chimpanzees imitate rationally, i.e. they choose whether or not to copy a demonstrator's actions. They concluded that the chimpanzees in the study had some understanding of the rationality of the demonstrator's actions, which contributed to their choice of whether or not to imitate the demonstrator. In another study, Carrasco et al. (2009) assessed an enculturated chimpanzee and found they were skilled at imitating gestures more than with objects. These skills suggest a theory of mind could have been cultivated alongside, and because of, correlated linguistic abilities, and the social interaction and training that developed both.

In terms of false belief, apes have been unsuccessful when administered non-verbal versions of false belief tests (Kaminski et al. 2008). In Krachun et al. (2009), two non-verbal false belief tasks were run: in one, a participant and a human competitor watched an experimenter hide a reward in one of two containers. When the human competitor left the room, the experimenter switched the containers. In an alternate version, the experimenter switched the containers when the human competitor was still in the room, but their back was turned. Human children that participated in this test were able to pass successfully, showing false belief understanding; however, none of the chimpanzee or bonobo participants were successful. The authors noted that apes looked more at the unchosen container, possibly implying that they were confused at the outcome. Looking times are often used as a measure of understanding in non-verbal infants, to test whether an event differed from their expectation.

In an experiment with monkeys, Marticorena et al. (2011) tested the false belief abilities of macaques, using an experimental design that had been used in prelinguistic children and involved looking times. They found that the monkeys did not appear to make predictions about the human presenters during the tasks, and that the results supported that the macaques did not possess a theory of mind.

Recently studies have used eye-tracking experiments to administer non-linguistic false belief tasks to apes. Krupenya et al. (2016) showed that orangutans, chimpanzees and bonobos watching videos looked first to an area where they anticipated a human actor would likely look for an object. However, out of the eight bonobos, fourteen chimpanzees and seven orangutans, only eight primates looked first at the target place in both experiments, and six primates did not look at the target in either experiment. The rest of the apes (more than half) looked at the target in one of the two experiments. Overall, there were more instances of the apes looking at the target instead of the distractor, or looking at neither. The researchers highlighted a possibility in the experiments, that "apes could solve the task by relying on a rule that agents search for things where they last saw them" (p. 114). Still, the researchers made an important contribution to the observation of behaviour in apes in false belief situations, and suggested their research supports the idea that apes can hold false beliefs. The apes included in their study were named, however their levels of enculturation, age or training were not listed. Looking at the data of the twenty-nine apes and their results against these variables would be interesting to assess the way in which language or other training played a role in their false belief success.

In an experiment concerning non-primates, bird species have been the subject of some study in relation to their visual perspective taking abilities. Emery and Clayton (2006) showed that scrub-jays re-cache food when they anticipate having been seen. Bugnyar et al. (2016) also showed that ravens took into consideration the visual access of others and their caches. Visual perspective taking is an important component that scaffolds the acquisition of theory of mind, as discussed in Section 2.4.

Given the varying levels of enculturation and language training amongst species in the above and other studies, and how those variables are usually not noted in the studies, a correlation between language and theory of mind in non-human species cannot be reached with the current data. Interestingly, Buttelmann et al. (2007) note that apes raised by conspecifics emulate (copy the end result of an action), and that enculturated apes are able to copy the actions of others in an imitative way. It could be hypothesised that language trained apes might out-perform other individuals when given false belief, appearance/reality, or mental state judgement tasks, as enculturated apes will be interacting with linguistic experimenters, and in some cases being trained in tasks that require linguistic training. Buttelmann et al. (2007) write that, "enculturated chimpanzees thus have some understanding of the rationality of others' intentional actions, and use this understanding when imitating others." Similarly, Tomasello et al. (1993) conclude from their study that, "Results showed that in immediate imitation the mother-reared chimpanzees were much poorer imitators than the enculturated chimpanzees and the human children, who did not differ from one another." While there are many developmental studies that show the link between theory of mind and language, the available evidence from non-human animal studies is needed. While some evidence is implied, it is however lacking.

2.7 Co-evolution of Language and Theory of Mind

There are researchers who posit a co-evolution of language and theory of mind (Corballis 2012, Malle 2002, Worden 1998, Smith 1996, Shatz 1994), because of their close developmental and functional relationship: "There is reason to believe that language and theory of mind have co-evolved, given their close relation in development and their tight connection in social behavior." (Malle 2002, p. 265), which Malle gives as either an adaptive demand, such as for improving cognition or memory, or an external demand that influenced the escalation of both abilities together in a sort of 'arms race', perhaps in reaction to the need for greater social coordination. These two possibilities are not mutually exclusive, and more importantly, are more parsimonious that they were one and the same adaptive pressure.

These discussions about the evolution of these capacities are necessarily informed and understood based on research of modern humans. This presents a problem to some researchers, who struggle to accept the homology argument that we can study modern minds to understand those from the past. Our cognitive evolution, like our physical evolution, is continuous with ancestral hominids in that there is a common origin, related through unbroken common descent. Our cognition has emerged from these ancestral forms which Palaeolithic hominins also possessed; as such, its organization should not be expected to be built upon radically different organizational features without supporting evidence to lead us to that conclusion.

After all, it is their organisation that through evolutionary changes led to our own. They are homologous constructions, not analogous, where similar forms have evolved for the same purposes out of different origins. That is why there is legitimacy in looking at the cognition of modern humans and hypothesizing backwards to the cognition of Palaeolithic hominins, who are only separated from modern humans by a few million years at most. We are more similar than we are different, and treating our species as distinct beyond cognitive comparison is to eject our species from the comparisons allowed for between other species and their ancestors. I extend the argument by Wynn and Coolidge, in considering Neanderthals and modern humans, to modern humans and ancestral hominins in general, (2012, p. 1), that:

"Neanderthals were so similar to us anatomically and genetically that we believe the default position – the null hypothesis [...] – should be that Neanderthals were no different. We must then present sound arguments for any differences we propose."

This proposal is not to argue that Palaeolithic hominins were cognitively equivalent to moderns, or even that that is likely, but simply that if cognitive relationships and features in the development in modern humans are found, and it is anticipated that Palaeolithic hominins shared those features too, that they would too share similar developmental relationships. The alternative to modern and past hominins having the same organizational principles between cognitive features would be that, in order to have such a disjoint, the slate was once wiped clean and then developed anew (and over just a couple million years or less). Our focus in this thesis is particularly on the Lower Palaeolithic hominins, such as *Homo erectus* and *Homo heidelbergensis*: species where there is less support for sophisticated cognitive and linguistic abilities due to a fragmented material record. However, it is still considered prudent here to believe that the neural/functional organization of these abilities, whatever their level, would be similar in regards to their relationship due to their common origin and evolutionary continuity with ourselves. While I make no claim in this section about the language or theory of mind abilities of Lower Palaeolithic hominins (at this point in this thesis, but see Chapter Six for an interpretation of what the evidence suggests), I do support that whatever their skill level in these areas were, they can be assumed to have the same close developmental relationship as exists in modern humans.

2.8 Modelling the Co-evolution of Theory of Mind and Language

On the basis of the evidence presented so far in this chapter, I will construct a co-evolutionary model for the development of language and theory of mind in human evolution. It begins with elements of adaptive pre-theory of mind capabilities, including the ability to attribute agency to animate objects, and acquiring the ability to follow the path of another's gaze towards interesting stimuli. Intention-goal understanding follows as an increase in intentional agency is realised. With these skills, the understanding that conspecifics have a different visual field and trajectory to one's own perspective, and second-level visual perspective taking develops. So far, these are abilities found in other species. These are necessary to the development of a key attribute in the development of theory of mind, which is joint attention. Joint attention makes point following and gestural communication possible, and allows for the possibility of fixing word reference. With a community able to reference to communicate thoughts, a lexicon (a mental inventory of words) can develop.

Having an increased inventory of communal references stimulates a pressure for the gestures to be learnable and transmittable. Through vertical (next generation) and horizontal (peer) transmission, these culturally learned references self-organise and maximise their contrastive and perceptive qualities (this happens in the self-organization of vowel inventories in languages, as well as other biological systems, Oudeyer 2005). In a spoken oral lexicon/call system, this would motivate the emergence of a phonological system (an inventory of meaningful contrastive sounds to construct words) to differentiate words. A phoneme is the smallest meaningful 'sound unit' in a language. In a language without phonemes (for example, a language where utterances resembled various animal calls, which are holophrastic and do not break down into constituent sounds or meanings), each and every call would have to be distinct enough from other calls in order to remembered and reproduced. By breaking up speech sounds into building blocks to construct contrastive and perceptually distinct sound sequences, the number of recognizable 'words' that can be constructed increases substantially. This sort of system would be motivated by the number of words that the community uses (in this use of words, 'utterance' may do better, for it is intended just to mean a standalone morpheme, which at a certain point in hominin linguistic evolution, might have been non-combinatorial, yet have no internal structure, but do the job of what in modern language might be done with a sentence made up of many linguistic elements).

Observing others use words to reference things (and not just objects, but situations and behaviours as well) reinforces the idea that others are mental agents, and a second-level intentionality is fostered. With the ability to hold a false belief and to consider other people's mental states, a true theory of mind has been achieved. Theory of mind opens up a shared mental landscape allowing for a cooperative mental world. An individual is no longer mentally 'alone', but is truly part of a community. The emotional affects possible with a community of mentalizing individuals provides the arena for the development of a lexicon to describe those mental states. Different syntactic elements such as complementation might have evolved in order to describe these mental states. In turn, exposure to others who are also describing others' mental states allows for individuals to appreciate that others also have theory of mind, promoting the ability to appreciate third-level intentionality and behave in a way in which they can predict how they will be perceived by others. This narrative can be illustrated in the schematic shown in Figure 2.4.

This model integrates information from child developmental acquisition of theory of mind and language, as well as evolutionary linguistic research into the self-organizing principles that hypothetically would have influenced the formation of complex language, including its phonemic, semantic and syntactic evolution. It creates the interpretive potential for cognitive and linguistic abilities if theory of mind could be indicated from Palaeolithic behaviour.

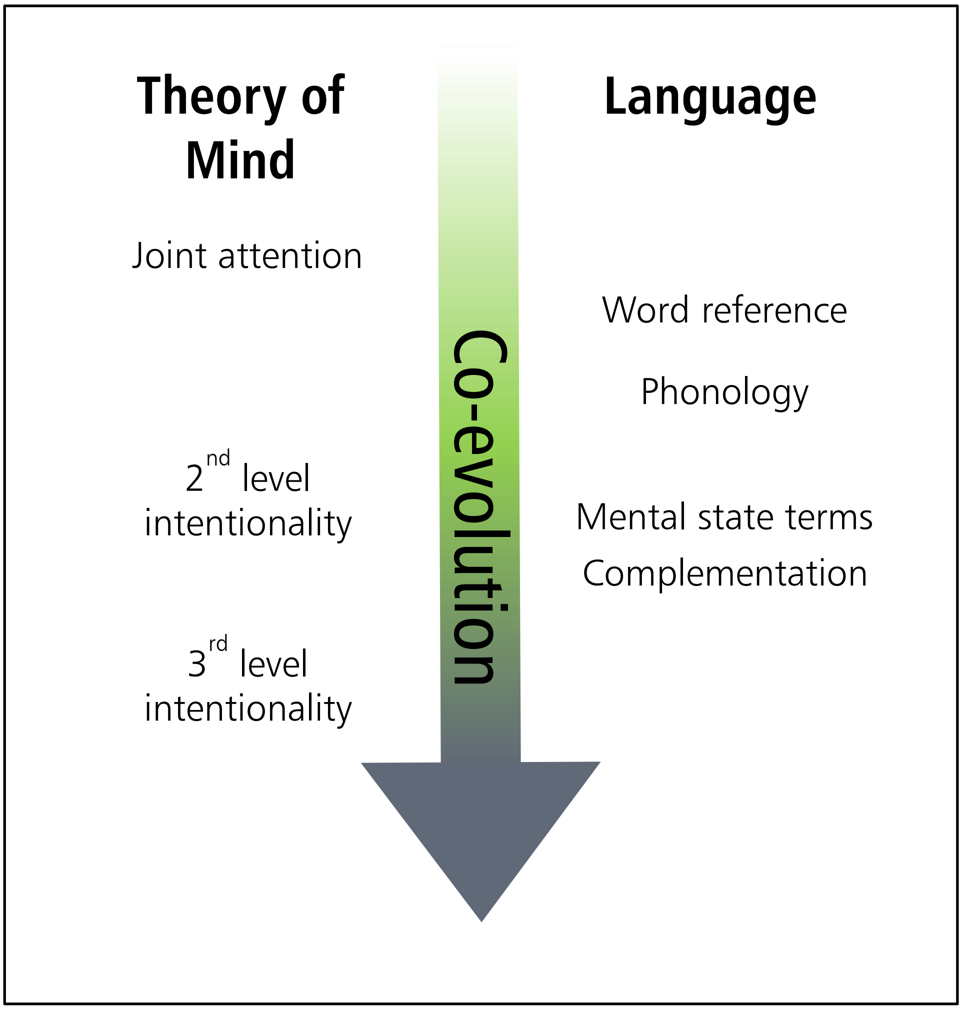


Figure 2.4 A model of theory of mind and language showing the order of its development in a hypothetical co-evolutionary narrative.

2.9 Summary

In this chapter I have examined at theory of mind: what it means, how it operates, the history of its study, and whether it holds a unique place in the modern human species. In addition, I have presented a selection of research that shows theory of mind's intimate developmental and functional relationship with language, that goes beyond just typically developing humans, but also children with developmental disorders such as autism, physical impairments such as hearing impairments, in addition to consideration of adults' abilities as well. The evidence suggests, although research is lacking, in looking at the correlated linguistic abilities in non-human primates, not to mention more diverse non-human species as a whole, that theory of mind and linguistic-like abilities also correlate. Future study may well be able to shed light on this likely relationship outside of the human species.

I have discussed why the information gathered from developmental studies should suggest that theory and mind and language shared a co-evolution and developed in conjunction with one another. I have proposed a narrative of what that evolution would look like, with theory of mind abilities influencing the development of language abilities, which allow for the further development of theory of mind and how that further influences the complexity of linguistic expression. As Malle (2002) suggests, this could indeed represent an arms race of the two abilities, in which they scaffold the complexity of the other, which reinforces the complexity they can attain. In this way, I have established that language and theory of mind are not only intimately related, but their relationship is such that this relationship can be used archaeologically. Language is notoriously difficult to study in evolution: it has been called 'hardest problem in science' (Christiansen and Kirby, 2003). Which means that theory of mind and its evolution might offer an easier opportunity to study the evolution of language indirectly. Theory of mind is a social behaviour, and social behaviours impact the archaeological record, leaving material indicators.

The next chapter will discuss social learning, the way knowledge is culturally transmitted through social interaction. Some methods of social learning necessarily involve theory of mind, and if these social learning methods can be identified in the archaeological record, we have an indirect indication of language ability possessed by hominins using this social learning method. With the relationship between language and theory of mind extended to social learning, we are one step closer to finding signs of Palaeolithic language use archaeologically.

Chapter Three.

Theory of Mind in Social Learning

3.1 Introduction

Having explored and established the co-evolutionary relationship between theory of mind and language, this thesis' aim of constructing a method to interpret language from Palaeolithic materials can continue. This will be done by looking for ways to identify not language, but theory of mind in the material record, which will be used as a proxy for language ability. The purpose of this chapter is to promote different social learning methods used to learn and transmit cultural information (some of which require theory of mind) as a link to detecting language ability in the Palaeolithic record. This chapter will therefore be establishing the second link in the inferential chain in Figure 1.1, that of social learning and theory of mind.

A description of social learning will first be examined, with examples of species that learn and share cultural information socially. Then three methods of social learning will be examined in more depth: emulation, imitation, and teaching. Imitation and teaching in particular are important for the purposes of this thesis as they require theory of mind to operate. Therefore, these types of social learning are an opportune behaviour to identify in the material record, in order to recognise the use of theory of mind in the creation of material, such as in stone tools. Emulation, though not requiring theory of mind, will have a focus because it is used to transmit material objects and tools in non-human species, though probably not in a cumulative manner (Tennie et al. 2009).

The chapter will then discuss how gaining information and knowledge through others affects the material culture humans create; there will be a review of some relevant cultural transmission studies, and recent experimental studies designed to explore cultural transmission of archaeological materials. Finally, the chapter will consider the chain of inference, where social learning implicates theory of mind ability, which implicates language ability. This approach to cognitive interpretation in the record is more appropriate than the dominant proxy for cognitive ability in the Palaeolithic record of symbolism (discussed further in Section 1.4).

3.2 Social Learning

Many species learn from the behaviour of others, a process which is called social learning (Heyes and Galef 1996). Social learning refers to receiving information "by observation of, or interaction with, another animal (typically a conspecific) or its products" (Heyes 1994, p. 207). This is as opposed to acquiring biological knowledge through genetic inheritance, or learning through individual/asocial learning.

Different types of social learning have been categorised and distinguished by researchers based on the different positions of the observer (the individual doing the learning), and the modeller or material the modeller produced (the individual replicating the behaviour or object, Heyes 1994). Lycett (2015) recognises four main types of social learning which he organises into a 'coarse taxonomy', which include stimulus enhancement, emulation, imitation, and teaching. This broadly encompasses all of the major types of social learning researchers typically recognise (but for a more comprehensive breakdown see Heyes, 1994). Emulation, imitation and teaching are the three methods focused on in this thesis, but stimulus enhancement too deserves a moment's attention, especially since it is so widespread in the animal kingdom. It does not require theory of mind like imitation and teaching do, and is considered to be less cognitively demanding than the other three categories.

The social learning mechanisms which most animals use do not involve theory of mind; however, theory of mind is an important and defining component of both imitation and teaching (Tomasello 1996, Straus et al. 2002). Emulation, a sophisticated form of social learning that does not involve theory of mind, is simulated in this thesis' experiment as it is used often both in human and non-human species, and has the ability to allow the persistence and spread of certain cultural practices. Most notably, it allows for the transmission of the creation of materials and use of tools, e.g. chimpanzee nut-cracking (Biro et al. 2003). Emulation, imitation and teaching will therefore receive the strongest focus in this thesis.

3.2.1 Stimulus Enhancement

Stimulus enhancement is the most commonly occurring social learning mechanism in the animal kingdom, and the other main categories (emulation, imitation and teaching) are relatively rare, especially imitation and teaching, which are often thought to be unique to human behaviour (Tennie et al. 2009). Stimulus enhancement is where a conspecific (usually, but not always) unintentionally draws attention to stimuli by its own interaction with it. This encourages interaction from a learner, which can result in the spread of cultural knowledge. A version of this type of learning is called location enhancement, which includes the same process, but it is the location that the learner is being drawn to that encourages the learning rather than the interaction with a stimulus. Heyes (1994, p. 214) describes the difference between the two as being a difference of exposure to stimuli, versus attention to stimuli, which she points out is a problematic distinction to distinguish empirically. For this thesis' discussions they will both be subsumed into the one category of stimulus enhancement.

A popular example of cultural knowledge being transmitted through stimulus enhancement is in the case of blue tits, a small Eurasian bird. During the earlier 20th century, along with the morning delivery of milk bottles, some blue tits learned to peck through the milk bottle tops to drink the milk on top. This cultural behaviour's spread was first documented by Fisher and Hinde (1949), and argued to have originated independently from several sites in the UK and Ireland and continuing to spread through social processes (Lefebvre, 1995). The rate of its spread was suggested to be the result of social learning, where Sherry and Galef (1984, p. 938) note, "It may be more appropriate to regard this, and possibly other instances of cultural transmission of learned behaviour observed in nature, as due in part to changes in the environment produced by those individuals introducing novel behaviours into a population." In an effort to characterise what type of learning was taking place with the blue tit example, Aplin et al. (2013) conducted a 'two-action and control' experiment, which meant to elucidate whether the innovations taking place were through social learning or independent innovation; in their experiment, two groups of wild-caught adult blue tits were exposed to a novel foraging task in one of two manners, and a third control group was not. If the exposed groups picked up the novel foraging tasks more than the control group did, it could be concluded that social learning had taken place and the birds had been transmitted that information. If emulation was taking place, then the innovations would match the novel foraging task shown to each group. The results of the study showed that around half of the time, birds in the experimental groups learned the foraging task, while those that were in the control group did not, supporting that social learning had taken place. They also found that juvenile females were much more likely to learn the novel task, and that dominance also played a role. They also interestingly showed that the blue tits were more likely to solve the foraging task in the same way the demonstrator did, meaning that they had also taken in specific details about the manner in which the modeller achieved the foraging task. That means that emulation (referred to as observational learning in this study) was also possibly taking place, and that the social learning mechanisms were perhaps not just a case of pure location enhancement.

Other forms of social learning that have spread cultural knowledge into species communities and likely driven by stimulus enhancement as a mechanism have been recorded in other gregarious animals as diverse as Japanese macaques (Kawai 1965), greylag geese (Fritz et al. 2000), black rats (Terkel 1996), humpback whales (Weinrich et al. 1992), cockroaches (Lihoreau and Rivault 2011), and bumblebees (Avargues-Weber and Chittka 2014). This diversity covers not only primate and wider mammalian species, but birds and insects as well. Humans, too, are included in this group as species who learn from the behaviours of our conspecifics who draw our attention to stimuli or to locations which results in learning, and this begins to happen from a very young age as infants (Horne et al. 2008). Humans, however, also draw upon a number of other social learning mechanisms in their repertoire.

3.2.2 Emulation

Emulation is an 'end-state focussed' copying behaviour. This means an individual replicates the product that resulted from another's behaviour (Huang and Charman 2005). It does not require an individual to theorise about the mental states of others, as the individual is not basing their behaviour on connecting another's actions and their goal, but simply the replication of an object or behaviour that they produced (Figure 3.1).

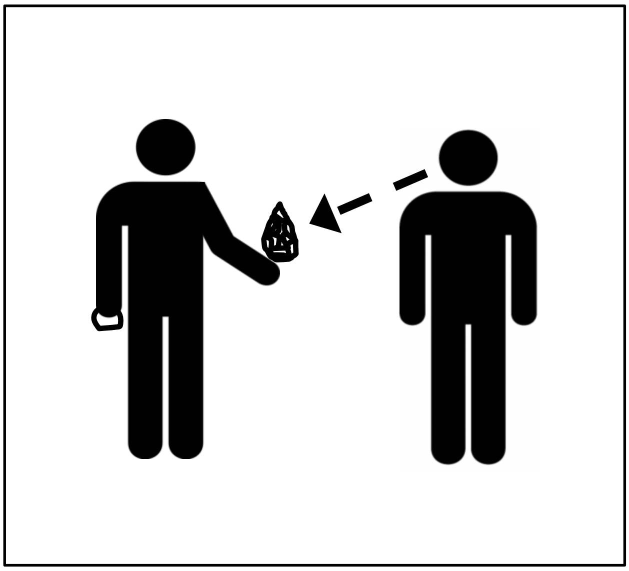


Figure 3.1. Emulation: the learner takes in information from the product to replicate its 'end-state'.

Tomasello et al. (1987) delineated emulation from imitation following a copying behaviour they noticed, when they observed chimpanzees learning a new behaviour by watching the function of a stick as a tool in an experimental condition. The chimpanzees learned about the affordances of the object visually, but did not replicate the processes they saw it being used in as in an imitative process, and so this was instead termed emulation (which similar behaviours have also been called observational learning, Heyes 1994). This type of copying behaviour was suggested then to limit the ability to transmit cultural knowledge due to its low-fidelity and lack of procedural knowledge.

One thing that Tennie et al. (2009) note as an important feature in the transmission of cultural knowledge is what they call the 'ratchet effect'; that is, accumulated modifications to culturally transmitted knowledge which stay stable in populations long enough for themselves to be modified, with relatively little knowledge loss or 'slippage', which then offers the opportunity for further modification and the chance to continue to be transmitted in a different form. This ratchet effect does not seem to be in effect in cases of emulation learning. For example, there are no accounts of it in chimpanzee learning, which the authors note tends to only include things that chimpanzees would naturally be able to figure out themselves independently. They suggest that for a ratcheting of cultural knowledge, there needs to be 1) inventiveness, and 2) faithful copying. While many species exhibit inventiveness, emulation does not allow as faithful of copying, as the process is not taken into consideration in the replication behaviour and therefore there is a considerable margin for deviating from the original process, which the learner does not have access to. As Tennie et al. (2009) describe it, it is as if the individual must reinvent the wheel each time. Humans, on the other hand, focus a great deal of attention on the process in order to replicate a behaviour, even to the point of including superfluous elements, called 'over-imitation' (Whiten et al. 2009). The stability this introduces in copying faithfulness could be what allows for a 'descent with modification' of cultural knowledge. This is supported in Wasielewski (2014), who conducted a study on modern participants who needed to complete a group task under different simulated social learning conditions. Results suggested that imitation was required for cumulative cultural evolution, where Tennie's 'ratchet effect' could take hold. In terms of the archaeological record and material cultural and its complexity, the ratchet effect has great implications for the behaviours and abilities involved in the persistence of certain material technologies in the record.

Much of the research regarding emulation has been conducted on chimpanzees, which have been long considered expert emulators, while humans have been considered the expert imitators (Tennie et al. 2010). Emulation is often contrasted with imitation, as imitation takes into account and replicates the process which resulted in the end-state, and therefore requires theory of mind to connect behaviours and intentions.

To test if emulative learning underlined chimpanzee social learning behaviour, "ghost experiments" have been constructed to create an impersonal scenario to be replicated, without the actions of a living agent. For example, this is simulated by manipulating the objects with fishing line (Hopper et al. 2007) so that it does not appear that they are being moved by a living agent. The hypothesis was that if chimpanzees are emulators, and they learned about the situation as a result of the object movement, then they should be able to learn from this type of model. However, the Hopper et al. study showed that chimpanzees only acquired the behaviour when another chimpanzee performed it, and not during the 'ghost' simulations, and so mapping movement onto another body may be an important part of the emulative learning process in cases when it involves manipulation of a tool.

Chimpanzee nut-cracking is a behaviour that is considered to be transmitted by emulation (Marshall-Pescini and Whiten 2008). It is a cultural behaviour in that it is learned (indeed some chimpanzee groups do not crack nuts at all, even if nuts are available to them) and there is variation in this behaviour between groups. It takes chimpanzees many years to learn to crack nuts in the wild, which requires them to balance a nut on an anvil and break it with proper force and precision with a stone or piece of wood. Typically young chimpanzees learn to crack nuts from their mother, and if a chimpanzee does not acquire the skill during a critical period (aged 3-5 years), it will not learn the skill (Biro et al. 2003). The qualities of different materials and the coordination of the different activities involved in nut-cracking each require much experimentation on the part of the chimpanzee to become an expert.

3.2.3 Imitation

Whiten et al. (2004) note that emulation and imitation overlap in ways that do not offer an easy dichotomy, especially when analysing animal behaviour both in controlled settings or in the natural environment. Here, however, I will treat imitation as a 'process-focussed' copying behaviour, where an individual replicates the behaviour (process) of another, because of a perceived goal/intent (Figure 3.2). This requires a level of interpretation and attribution of intent to a species, which can be difficult. Tomasello et al. (1987) identify three criteria to help recognise imitation learning in lab experiments: 1) the replicated behaviour should be novel, 2) it should reproduce the behaviour of the model, and 3) it should have the same final goal as the model (perceiving intent).

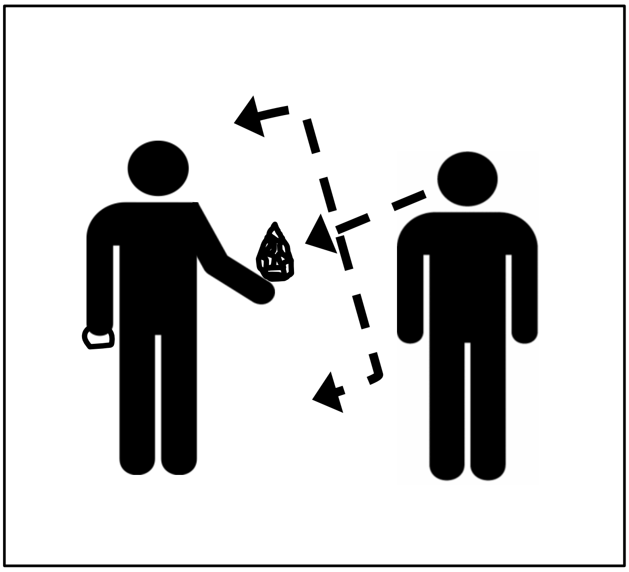


Figure 3.2. Imitation: the learner is focused on replicating the process that led to the end-state. The wider breadth of arrows (compared to emulation, illustrated in Figure 3.1) represents the attention being paid to the maker's bodily processes that created the product.

For this thesis' purposes, the perception of the goal and its connection to the process that resulted in it is important because of the cognitive mechanisms this behaviour requires; in order to imitate, an individual must have the ability to attribute intention to another individual's behaviour, linking it to an outcome which fuels the motivation for its replication. This is the narrow definition of imitation, as opposed to defining it as any direct copying behaviour, which in some instances is called mimicry (Tomasello 1996). Mimicry is simply replication of actions without a perceived goal, such as when a parrot mimics a human voice in a non-communicative way. Mimicry appears in a number of gregarious species, such as birds which learn their song from conspecifics. Imitation can be assumed instead of merely mimicry in cases where the behaviour signals investment or a likely function, such as in the case of the creation of a stone tool in the archaeological record, if it took steps where the cultural knowledge for its creation is such that it would not be learned individually.

Illustrating a case of imitation, in a study by Meltzoff (1988), 14-month-old participants were shown a box with a panel on the top surface. The experimenter demonstrated to participants a causally-irrelevant behaviour, by bending from the waist and touched the panel with their forehead, which illuminated a light bulb in the box. The majority of infants copied this novel behaviour when given the box, while the control group (who had not seen the experimenter's actions) did not, showing that it was indeed a novel behaviour. Meltzoff concludes that by 14 months, "imitation is well-enough articulated to play an important role in early learning and development." (p. 475). However, Range et al. (2007) calls young children's imitation inferential, selective imitation that does not involve theory of mind, and is more akin to mimicry. This, "does not require the attribution of mental states to others but relies simply on the evaluation of the observable facts: the action, the goal state, and the situational constraints." (Range et al. 2007, p. 868). This illustrates the difficulty in ascribing imitation in experimental settings to both human children and non-human species.

Imitative ability seems to increase as human children get older; McGuigan et al. (2007) studied three year old and five year old children and found that five year olds were even more likely than the three year olds to copy all of the actions that they saw, even when they could be seen to be irrelevant to a task. Children have therefore been referred to as "over-imitators" (Whiten et al. 2009): A number of 'artificial fruit' studies have been conducted by Whiten and colleagues (e.g. Whiten et al. 1996) showing chimpanzees performing imitative behaviours (both bodily and object movement copying). In the experiments, a special transparent box was constructed that in order to be open, needed to be manipulated in a certain way. Inside the box was placed a food reward that could be seen by the participant. Participants were then shown a method of acquiring the food reward contained in the box before allowing them to try. In certain conditions, the method for opening the box included superfluous actions, which results showed children were much more apt to copy.

It is not known whether imitative behaviour in chimpanzees is limited to those of enculturated chimpanzees, which may have acquired the ability through development of other cognitive abilities that bolstered skills such as understanding the intents of others, including communicative intents. Some experiments have been carried out on mother-reared chimpanzees as opposed to human-reared chimpanzees, but all chimpanzees have a certain level of human interaction and behavioural expectation. Whiten et al. (2004) note that, "the evidence for imitation of the highest fidelity has typically come from apes that have been raised with much interaction with humans." (p. 45-46), and "It is not likely that either monkeys or apes who have not had extensive human contact imitatively learn the instrumental behaviors of others." (Tomasello and Call 1997, p. 294). If it is the case that chimpanzee imitation only occurs with the socialising and intention-understanding that comes with an interactive, linguistic, socially complex environment, it lends strength to the hypothesis that social learning, theory of mind, and language are all intimately connected and mutually supportive abilities with relationships that exist not just in human species, and likely premodern humans as well.

As for examples of imitative behaviours in animals other than chimpanzees and humans, there have been instances recorded in a number of species (Voekl and Huber 2000, Zentall 2004), and contention surrounds the extent to which it should be considered imitation, and what is the mechanistic background to the actions. For example in Range et al. (2007), they showed that dogs were able to perform the same actions that a model dog demonstrated for them using their mouth to pull a rod, instead of their paw (a more intuitive action). They suggested that dog imitation is more akin to infant imitation than chimpanzee imitation, because chimpanzees are less inclined to imitate an irrelevant action.

Whiten and colleagues have challenged the notion that humans are the sole imitators and chimpanzees are mainly emulators (Whiten et al. 2009), instead arguing that "both chimpanzees and children possess a 'portfolio' of different social learning mechanisms, including both imitation and emulation, that are deployed selectively in different contexts." The authors also draws a distinction between 'can imitate' and 'does imitate' – they see chimpanzees as having the capacity to produce imitative behaviour in some instances (and this could be a product of their training and interaction with intention understanding beings), but that they prefer not to, or do not have the inclination to. In experiments, object movement seems to dominate as chimpanzees preferred focus. For example, Myowa-Yamakoshi and Matsuzawa (1999) noted that chimpanzees found it easier to replicate object-object actions. It seems illogical, though, that a species would struggle to observe and learn how to acquire a skill such as nut-cracking, which takes a great cognitive investment in time and effort and frustration, if this type of learning was in their repertoire but they did not use it and better connect the processes involved. The research in animal social learning runs parallel to animal theory of mind, in that gregarious animals seems to exhibit features, which may be accentuated in enculturated individuals, but events in the wild do not seem to exhibit these traits.

3.2.4 Teaching

In addition to this 'portfolio' of different modes of cultural transmission (and humans use all of stimulus enhancement, emulation, and imitation behaviours too), humans also teach. What is meant here by teaching is specifically, "an intentional activity to increase the knowledge (or understanding) of another" (Frye and Ziv 2004, p. 458). So while teaching does not describe the learner, as the others do, it describes the knowledge state of the modeller interacting with the learner. This intentional/cognitive definition used by Frye and Ziv is crucially different to the functional definition put forth by Caro and Hauser (1992, p. 153):

"An individual actor A can be said to teach if it modifies its behaviour only in the presence of a naïve observer, B, at some cost or at least without obtaining an immediate benefit for itself. A's behaviour therefore encourages or punishes B's behaviour, or provides B with experience, or sets an example for B. As a result, B acquires knowledge or learns a skill earlier in life or more rapidly or efficiently that it might otherwise do, or that it would not learn at all."

This definition includes activities that might be biologically motivated, such as a felid bringing back live prey for young to develop hunting skills. While both definitions include the active participation of instructors, Frye and Ziv's definition highlight it as a purposeful, goal directed activity of knowledge transfer (Strauss et al. 2002). In order to teach, the teacher must believe they can impute knowledge and that it can be received and understood by another individual. Intentional teaching, then, can only occur if the teacher has the ability to theorise that the student has the ability to imitate them. The teacher must be able to conceive that others think about what others think about. This requires a higher capacity for theory of mind than imitation does. Teaching not only requires an individual to have beliefs about the knowledge states of another (Strauss et al. 2002; Kruger and Tomasello 1996), but to appreciate that another will theorise about what they themselves think. In other words, in teaching, *a* thinks that *b* will recognise what *a* knows (and is intentionally trying to show them). Thus I make the argument that teaching requires third-level intentionality (Figure 3.3).

While the literature does not advocate more complex theory of mind for the ability to teach (most of the focus is on the acquisition of theory of mind and children ages 3-5, rather than more complex applications of mental states), it is unanimous in the involvement and important role that theory of mind does have in being able to teach. Ziv et al. (2016) tested seventy-five 3-5 year olds in a number of theory of mind and 'understanding-of-teaching' tasks. A subset of these children were then randomly chosen to teach a board game to their peers. The three year olds taught using demonstration, but the four and five year olds reacted to the changing knowledge level of the learners, showing their attention being paid to the knowledge states of the learners. The authors concluded that their development of theory of mind was closely related to the ability to understand their concept of teaching.

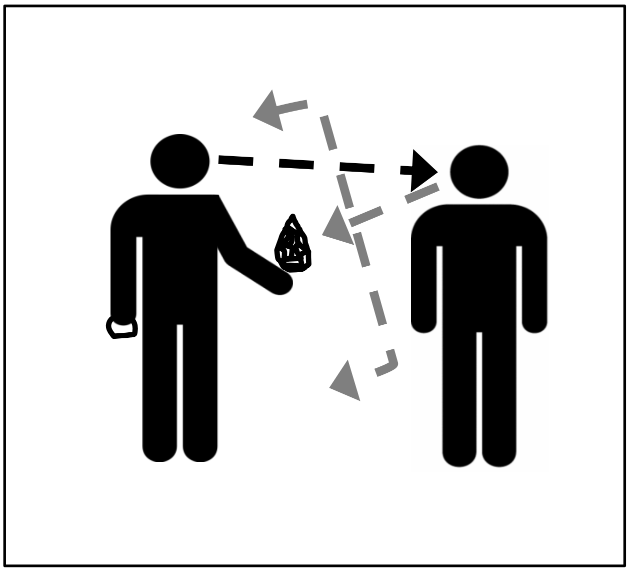


Figure 3.3. Teaching: the teacher is intentionally conferring knowledge to the learner, because they assume the learner's capacity to imitate and receive the knowledge. This requires third-level intentionality on the part of the teacher.

Amon recent claims of teaching in non-human animals, Musgrave et al. (2016) document the behaviour of chimpanzees while termite-fishing, where mothers allow young chimpanzees to take their termite probes from them. The authors note how this is a costly activity as it takes away from their feeding time, and by providing them with the tool probes, they are helping the young chimpanzees to learn about the activity. They hold that this behaviour fits the criteria for teaching because 1) it happens in the presence of the learner, 2) the behaviour is costly to the teacher, and 3) it improves the performance of the learner. However, there is no evidence that the knowledge being conferred is 'intentional', nor is the behaviour any different than the facilitation of learning caused by chimpanzee mother who allow young chimpanzees to take tools and nuts from their hands, interrupting nutcracking. Rather than an example of teaching, the videos provided in the supplementary information show an example of how chimpanzees are tolerant of their young, and this facilitates their learning of complex behaviours.

The roles in the attribution of mental states differentiates the theory of mind involvement of social learning methods used in humans and non-human species. While the mechanisms and underlying principles are hard to tease apart in lab experiments and in the natural environment, the impact these mechanisms have on the transfer of knowledge and the replication of materials and behaviours is important. Further comparative studies with non-human species, especially with a better understanding of the mechanisms behind imitation, and how imitation-like behaviours are produced by animals who do not possess a theory of mind, are crucial to the construction of the hypothesis in this thesis. What is clear is that material culture and the preservation of modified cultural traditions in non-human species is rare. The 'ratchet effect' that looks to be crucial for the transfer of complex technological information such as found in idiosyncratic chipped stone technology (anything beyond the passive percussion or bipolar techniques resembling chimpanzee nut-cracking or early hominin flake and core technology), surely require social learning mechanisms afforded by imitative learning in order to behaviourally persist. Further understanding the different social learning mechanisms' impact on the creation of material culture will allow a more confident interpretation of different Palaeolithic stone tool assemblages, and the cognitive mechanisms that they allude to. The crude separation of emulative, imitative, and teaching strategies in the transfer of material culture is likely a reductive one, where strict boundaries do not exist, but are instead gradient divisions. But understanding their unique characteristic signatures and morphological expressions in the material record will undoubtedly highlight some of the cognitive necessities involved in the teaching and learning abilities that Palaeolithic ancestors used in the cultural transmission of their gathered technological expertise.

3.3 The Impact of Social Learning on Material Culture

Archaeologists have analysed variation in materials to study the past as long as the discipline of archaeology has existed (e.g. Evans 1861). Indeed, culture historical approaches define the earliest explanations of material variation, and were the dominant approach in the early 20th century and before. The way that culture, including material culture, changes over time is described by cultural evolutionary theory (Cavalli-Sforza and Feldman 1981, Boyd and Richerson 1985) which takes Darwinian principles of variation, inheritance, and selection to produce change over time. Biologists have long used these principles to understand species change, but cultural evolution too is guided by these principles; not in an analogous way, but in evolution's truest sense (Mesoudi et al. 2006, Lycett 2010). Studies of these cultural evolutionary processes, however, are much less numerous. Laboratory experiments have been productive in studying the processes underlying material variation (including those that use computer modelling to predict null hypotheses of cultural change), which can compare results with the material record to find evidence of different processes (e.g. Eerkens and Lipo 2005).

3.3.1 Experimental Cultural Transmission Studies

A number of studies have looked at cultural transmission with a view to understanding technological change in the archaeological record, and especially recently, this has become a popular topic of research (e.g. Stark et al. 2008, Ellen et al. 2013, Mesoudi and Aoki 2015). Within the research, a handful of experimental approaches have been conducted in controlled circumstances with the production of stone tools, to elucidate more clearly the variables at play and how they would materially manifest themselves in the cultural evolution of lithic technologies.

In the first of these types of experiments, Mesoudi and O'Brien (2008) designed an experiment to simulate the cultural transmission of New World projectile technology, exploring two of the modes outlined in Boyd and Richerson (1985), a paper which discusses a number of biases that happen in cultural transmission. Boyd and Richerson discuss "guided variation", where a cultural trait is inherited through social learning, but then modified through the individual's own trial and error, introducing diversity. "Conformist bias" is where a variant of a cultural trait that is the most popular is the most likely to be adopted in a population, reducing diversity. Finally, "indirect bias" is where a variant of a cultural trait that is used by a high status individual is more likely to be adopted by others. Mesoudi and O'Brien (2008) explored two of these biases, indirect bias and guided variation, in an experiment simulating these cultural transmission processes in the laboratory with a simple video game. The video game was designed to simulate hypotheses about Great Basin projectile points made in Bettinger and Eerkens (1999). Eighty-one Participants used virtual projectile points in three different virtual hunting environments. In the first part of the experiment, which represented oblique cultural transmission and tested indirect bias, participants were shown a previous groups' point designs and success rates, and asked to copy one of the point designs. The hypothesis predicted that the most successful participants' projectile points would be chosen and copied more often, and therefore the points would converge. Results showed that this was the case, and in the majority of the cases participants chose to copy the most successful point design, showing oblique cultural transmission. In the second part of the experiment, participants were able to modify their own virtual projectile points in order to experiment with their success rate, but with no access to others' points (which represented an individual learning stage and tested guided variation). With this individual learning, the hypothesis was that the point designs would diverge, possibly through a process of drift or adapting to the different virtual environments, which had different parameters and optimal variables. Again, results were consistent with the hypothesis, and the point variables diverged, which they attributed to the adaptive virtual landscapes. Finally, in the third part of the experiment, participants were able to modify their projectile points and were also able to copy the members of their group (if they wished) and see their success rates, in an indirectly biased horizontal transmission. This allowed individuals to copy the most successful in their group, and the results showed they did. Groups had different participant numbers for comparison: there were groups with between 2 and 6 members, as well as individuals who participated without belonging to a group. One of the interesting impacts of group cultural transmission they found was that within-group variation fell when there was cultural transmission, which alludes to the hypothesis in this thesis, that low morphological variation could be an indicator of sophisticated cultural transmission methods. Finally, participants who were in groups performed better than those that were working individually, indicating that the indirectly biased cultural transmission they used was adaptive in finding more optimal projectile point designs.

Kempe et al. (2012) also conducted an experiment simulating cultural mutation in an archaeological context. With 200 participants, they conducted 20 different transmission chains where participants had to, as faithfully as they could, resize a handaxe to match the size of a model on an iPad. Their predictions followed from Weber's Law, which is a psychophysical description of how perceptual errors increase according to the magnitude of the attribute belonging to the object that is being perceived. Therefore they predicted that their transmission chains would retain the same mean size over time, while increasing in variance through the generations. This is a key principle for this thesis' experiment, as the same culture mutation principle will create copy error in those participants' stone tools as well (although the experiment presenting in Chapter 4 has only one 'generation'), although the different social learning methods used in each group are influencing the amount of variance being produced throughout.

The next step in Kempe et al. (2012)'s study was to test their results against the Acheulean Biface Database (Marshall et al. 2002). Length and breadth measurements were available for 2601 complete handaxes from a number of sites in five different countries. The handaxes in the dataset strongly deviated from the predictions in their model of means and variances, but they suggested a number of reasons why this may be: for one, handaxe transmission was unlikely to have passed in independent lineages. The database was also limited in its temporal and spatial coverage. Functional limitations might have also created a pressure on size regulation, such as fitting in the hand. And finally, shape might not have been a goal in the social transmission, and other factors were being copied that reinforced the size regularity. One suggestion for the deviation for their model that they did not suggest was that in their experiments, each generation only had one handaxe to copy from, instead of the accumulation of previous generations' handaxes as models. The multiple generations creating the model would mean that the average and the variance would be a function of previous generations as well, not just the model generation, influencing the rate of copy error.

Many experimenters have considered knapping in experimental settings as dangerous, as well as difficult to organise, and so Kempe et al. (2012), as well as others (e.g. Schillinger et al. 2015) chose to simulate the replicative process of copying a stone tool by resizing a handaxe on a touch screen to avoid healthy and safety and logistic issues. As the task was then relatively simple to execute, they could be sure that they were measuring just perceptual error rather than manufacture error or other variables introduced from the experimental design.

3.3.2 Knapping Experiments Exploring Social Learning

Schick and Toth (1993) suggested that an interesting experiment would be to train modern humans, verbally and non-verbally, in making stone tools, and to analyse their differences. They suggested this might inform us about the complexity of communicative abilities of pre-modern humans. Until recently though, only four studies have experimentally analysed the effects of different teaching methods on the morphology of stone tools (Ohnuma et al. 1997, Putt et al. 2014, Morgan et al. 2015, Schillinger et al. 2015).

Focussing on the difference between silent instruction and verbal instruction in the production of Levallois technology, Ohnuma et al. (1997) taught 20 participants (two groups of 10) Levallois technique (where a flake is removed from a predetermined core), using a siliceous shale material found on the riverbank where the experiment took place. Participants underwent training and knapping for a total of eight hours over four sessions. The experimenters predicted that the verbal group would acquire the technique faster than the non-verbal group, which would be taught the same principles, but through gesture. Material that the participants produced was assessed by two measurements, 1) how long it took them to be able to prepare a core, and 2) how long it took them to successfully detach the final flake. Nine out of ten (verbal) and eight out of ten (non-verbal) participants were able to prepare a core respectively, and six people in each group successfully detached a flake. The variances of the rate of acquiring the technique (successful preparation of a core) were not significantly different, but variances in the success of flake production (detaching the desired flake from the prepared core) did differ statistically. However, the means for both of these measurements did not differ between groups. The authors therefore concluded that, "spoken language was not indispensable for Levallois flake production in the Stone Age, and that this unique tool-making method belonged to a different level of subsistence activity from that which necessitated language." (p. 167).

A number of issues with this study include that the two groups were conducted in different areas, and with different raw material types. The verbal group which was conducted first used up the raw material, and the site had to be moved to conduct the non-verbal one. The verbal group therefore had been conducted on a courser grained material. The raw material, being a siliceous shale, also likely possessed preferential planes of fracture which would not be ideal for a controlled material. Another issue is that the verbal group had all previously taken classes in lithic technology, while the non-verbal group had not. Both of these issues had the potential to severely affect the study outcome, and therefore the results should not be taken with confidence.

Their observation that the non-verbal group was so adept at acquiring the preferential flaking technique is interesting, considering how researchers and knappers often describe Levallois technique as being difficult to acquire, and that it is generally a skill achieved only after much training; therefore, it is interesting that the verbal explanation of the techniques were not required. It means participants were able to pick up enough of the technique through observable cues, although their attention was undoubtedly intentionally being drawn to these cues by the teacher. Gestural communication of the type used, along with intentional teaching, would have been rife with theory of mind-laden interactions among the group and between participants. The teacher would have been intending that the students copy them – and therefore, third level theory of mind would have been used in order to create the environment necessary to culturally transmit the information both to the silent and the verbal groups. Through the literature in this thesis, it can be seen that complex theory of mind does not develop without the scaffolding provided by language and a linguistic social environment. So while 'language' is not necessary for the specific situation of transmitting Levallois technique, it certainly would have been necessary for developing the ability to transmit in that way with gesture. Because of this, then the conclusion of the researchers that language was not necessary to acquire Levallois technique is only partly true at best, and misleading at worst; language-less individuals would not be able to acquire the technique – only linguistic individuals who were all for some reason choosing not to communicate verbally.

In another study of the effects of verbal versus non-verbal teaching on the acquisition of stone tool technologies, Putt et al. (2014) conducted a similar but more controlled study in which a number of handaxes were created experimentally by twenty-four novice knappers taught under a silent or verbal condition. This study, similar to the Ohnuma et al. (1997) study, reported no significant differences pertaining to the quality or symmetry of the handaxes in the two different groups, but an analysis of the flake debitage revealed that the verbal group were making larger, more invasive flakes with larger platforms. They suggest that with their findings, "some inferences could be made about the linguistic behaviours of earlier hominin species." This is a much more cautious suggestion than conclusions drawn in Ohnuma et al. (1997), but it still makes the heavy allusion that the non-verbal or silent group was a simulation for a language-less, non-linguistic group, which it was not. The inclusion of gesture, and of shared communal signals in the non-verbal group is behaviour highly dependent on not just theory of mind, but presumptions and interactions built on a shared culture, which they are applying in the experimental task and have not been controlled for. Their interactions are of a more complex kind that has been scaffolded and realised with the socialisation that came with an environment rich in mental interaction, dependent on rich linguistic interaction.

The Putt et al. (2014) article, however, avoided many of the issues found with the Ohnuma et al. (1997) study. It was more controlled (although they also experienced an issue with consistency of raw material), included a much higher number of data for analysis (334 stone tools in total), and they chose to replicate handaxes rather than the Levallois technique, which lended it some benefits in the available analysis that could be conducted due to the imposed shape. Putt et al. employed a type of outline analysis called Elliptical Fourier Analysis (EFA), a type of geometric morphometric methodology which analyses the sine/cosine transformation of homologous outlines. The 2D closed outline of the planar view of the handaxes was converted into coordinate data from digital images using specialised software. The data could then be statistically analysed. EFA is a useful tool to compare ovoid shapes, depending on the number of harmonics employed in the analysis (higher harmonics mean more detail is retained in the outlines), although the number of harmonics used in the analysis was not noted. A Principle Component Analysis (an ordination method that looks for the largest sources of variation in the data) was conducted and provided a list of the features that varied the most within the stone tool assemblages. They found that the groups did not differ significantly from each other. Among other types of analyses were a flip-test for symmetry, as well as an independent rating for skill and quality from an experienced analyst. They did find that the debitage created by the two groups was different, with the verbal group producing larger flakes.

Confusingly in the Putt et al. (2014) study, the authors concluded that the differences in the two groups' learning strategies could be described as 'emulation' in the non-verbal group, and 'over-imitation' in the verbal group. As the non-verbal group undoubtedly were copying the processes of the model knapper in order to replicate similar tools, emulation does not sound like an apt description of the behaviour, even if the participants' resulting tools deviated from the more skilled models'. Imitation is a much more likely social learning mechanism at play because of the process replication, unless it was noted that many of the participants were seen to be attempting to make handaxes through means other than those that were being taught. In the verbal group, it is not clear what actions are being described as over-imitative. Over imitation usually refers to replicating behaviour even though they are superfluous to the task at hand (e.g. Whiten et al. 2009). The Putt et al. (2014) article implies that the experiment is structured to elucidate the relationship between language and cultural transmission in stone tool technologies in a way that may reflect some information regarding past hominin species and their technological transference. However, this article and the Ohnuma et al. (1997) article are not controlling for cognitive factors through their different groups, but rather behavioural factors, and the results of their studies reflect this, and are not to be taken to support evidence of cognitive necessities regarding cultural transmission and language ability.

In another experiment following up the link between cultural transmission of stone tool technology, and the relationship between teaching and language, Morgan et al. (2015) conducted the first knapping study of different learning techniques paired with an iterated learning methodology, and also expanded the study to include other social learning techniques. Their focus was on core and flake technique, and questions specifically about the Oldowan lithic industry. 184 participants learned to detach flakes by learning from another participant (who had learned from another participant, and so on) under five different learning conditions, as opposed to just the non-verbal versus verbal learning condition of the previous two studies discussed. The transmission chains were performed with two short (five participants) and two long (ten participants), with the first tutor being one of two experienced knappers. The participants were also motivated to transmit their technique effectively as they were paid for their time.

The different learning conditions were termed 1) Reverse Engineering, 2) Imitation/Emulation, 3) Basic Teaching, 4) Gestural Teaching, and 5) Verbal Teaching. Reverse Engineering resembled Group 1 (Emulation), which will be presented in this thesis' experiment in Chapter 4. The participants were shown the flakes that the tutor made, but were not able to witness the process of them being made. They called this group 'reverse engineering' rather than emulation, although it simulated a replication process where only the end product provides information about how to attain the goal, and the individual must find the process out for themselves, which by definition is an emulative activity (Huang and Charman 2005). The next group was then referred to as 'Imitation/Emulation', as the participants observed the tutor making the flakes. This is analogous to Group 2, Imitation that will be introduced in Chapter 4, although Group 2 (Imitation) differs in that it was delivered by video to prevent accidental slowing or emphasis of technique on the part of the teacher. The group had visual access to the processes to replicate the tool, and could use imitative behaviour. Including the term 'emulation' in the label could be misleading, as emulative learning could equally be applied across all group situations, including the teaching ones, so it was probably not necessary to mention. Their third group was called Basic Teaching, as it included the allowance of the tutor to shape the pupils hands, and also slow their actions or try to get the pupil to better see an action. The behaviours associated with a tutor slowing actions and manipulating hands would show the intention that the other individual make the tool. This would require theory of mind at level of intentionality two at the least, and closely resembles teaching at level of intentionality three, and so it would be a surprising behaviour if evidenced in non-human primates that were not brought up in a theory of mind-rich environment. The fourth group was called Gestural Teaching, and resembled not only Group 3 (Silent) that will be introduced in this thesis' experiment in Chapter 4, but both non-verbal groups in Putt et al. (2014) and Ohnuma et al. (1997). This group could interact with the tutor and use gestural communication, but silently. Finally the 5^{th} group was called Verbal Teaching, which resembled Group 4 (Verbal) in the teaching experiment as well as the two verbal groups in Putt et al. (2014) and Ohnuma et al. (1997). In this group the tutor and pupils were able to interact with verbal communication as well.

They found that while low-fidelity transmission methods such as emulation and imitation were not very successful in transmitting skill in the transmission chains, transmission through intentional teaching, especially when aided by language, were much more successful. The quality of the flakes only improved with language and gesture. They found that imitation/emulation did not result in an increase in viable (over 2cm) flakes, but verbal teaching had a clear increase. Through all the measures, they found that teaching, but in particular verbal teaching, greatly increased the success of measures. The authors suggested their findings are in line with studies that suggest that Oldowan flake knapping selected for increasing complexity in language and teaching (e.g. Wynn et al. 2011). They made the interesting conclusion that the imitation/emulation group made a minimal enhancement to the participant's success compared to reverse engineering. This could suggest, then, that the skills and behaviour involved in flake and core technologies do not require the more complex learning mechanisms that imitation provides, in order to acquire the technological skill. Oldowan technologies might have been able to spread and persist in a species with emulative behavioural abilities alone. Their conclusions were that gene-culture co-evolution would have provided a relationship between tool use and social transmission.

While the Morgan et al. (2015) paper referred to their finding as having implications for the evolution of language and teaching, they did not delve into the cognitive significance of the different learning traditions in terms of their necessary theory of mind abilities. One of the authors of the paper was Andrew Whiten, a leader in the research on social transmission and theory of mind in chimpanzees and developing children for decades, so it is a reasonable next step. The experiment presented in this paper uses the important studies conducted by Whiten and other colleagues, in conjunction with experiments such as these, to make the conclusions that can be drawn from experiments with modern participants reach farther.

A final experiment exploring the impact of different social learning techniques on resulting material culture happens with Schillinger et al. (2015). While not a study with language as a variable, they focussed on emulation and imitation more concretely. They conducted a study where sixty participants carved 3D handaxe shapes out of hard garden foam with a plastic knife, rather than using flint as in the other studies. While the foam was easier to source and safer to manage, and did not require a teacher trained in knapping, it did mean that an analysis of the 'flakes' was unable to be conducted. However, because of the standardised starting shape and sizes of the beginning piece, statistical analysis of the shape of the end products could be compared in a method not available to the flake and core behaviour of the Ohnuma et al. (1997) study, the non-standardised flint handaxe blanks of the Putt et al. (2014), or the non-standardised flint cores of the Morgan et al. (2015) study. This was a key advantage that the Schillinger et al. (2015) study had over the other three knapping experiments in that it allowed a much more in depth analysis of the material output.

In the study, one group had a model foam shape (in the shape of a handaxe) to replicate, resembling the set up for Group 1 (Emulation) in this thesis' study, as well as the Reverse Engineering group in Morgan et al. (2015). Another group viewed the foam shape being made by a model on a video, similar to Group 2 (Imitation) in this thesis' experiment as well as the emulation/imitation group in Morgan et al. (2015). The imitation group also had access to the model foam shape for reference.

In contrast to Morgan et al. (2015), they concluded from the results that the imitative learning condition significantly reduced the amount of copy error in the material produced; therefore, imitation learning could provide cultural traditions with an increased likelihood of transmitting and persisting. This hypothesis is in line with other studies that promote imitation learning as having the potential to create persisting and possibly cumulative cultural traditions, as opposed to emulative behaviours (Tennie et al. 2010). Lycett (2015) notes the importance of material culture on emulation learning, in that it creates the opportunity to replicate an end-state in the first place. This is interesting, as emulation only seems to exist in species that have cultural material; which makes sense, as material culture is the product of a behaviour, and emulation is the copying of a product of behaviour, which in many cases this requires a tool. This allows for real cultural evolution, in which material culture is truly descent with modification; Lycett (2015) describes the elements of Darwinian evolution in which social learning is the mechanism for descent, variation is introduced through innovation or copy error, and selection happens through biases in persistence. With high fidelity cultural transmission, real cultural evolution can take place. Schillinger et al.'s study, however, does not discuss the relevant connections that can be made from cultural transmission to cognitive and linguistic abilities as in the other papers, at least in regards to language. A productive extension of this study would be to consider the cognitive implications that the controlled cultural learning simulations imply.

These four studies have investigated how novice knappers learn to make stone tools under different learning conditions. Understanding what morphometric variables distinguish lithic assemblages made by individuals and communities who learned through different modes of cultural transmission is an important step in recognizing imitation and teaching in the archaeological record. However, none of these studies have connected the mode of cultural transmission to required cognitive levels such as with theory of mind ability. The argument presented here, that modes of cultural transmission imply theory of mind ability which imply specific linguistic prerequisites, is an original contribution of this thesis not found in previous knapping studies. This experimental study includes elements of the methodologies found in some of these previous studies, but does so in a way that has far reaching implications for understanding the effects of different modes of cultural transmission on material, so that these cultural modes might be identified in the archaeological record; namely, the introduction of theory of mind as a proxy, explicitly linking the mode of cultural transmission to the level of language ability present in those hominins.

Three of the four knapping studies outlined here were published during the course of the preparation of this thesis, which attests to the growing interest in applications of experiments of transmitting lithic technology and their implications for cultural transmission in human evolution. Issues such as raw material, standardisation of material, and simulating the behaviours involved in the cultural transmission, are all variables that continue to progress. The four studies here contributed a great deal to the inspiration for the experimental design and resulting analysis in the experiment outlined in Chapter 4, even when some of the inspiration had to be made post-hoc due to the timing of the publications of these recent papers. But what they all contribute to is the offer of an alternative way to access important behavioural and cognitive information about human evolution. Table 3.1 summarises the attributes of the four above noted studies alongside the present study detailed in Chapter 4, for easy comparison of method.

3.4 Conclusion

This chapter has been about social learning, its role in cultural transmission, and its relationship with theory of mind. It covered the utility of experimental approaches to understanding how cultural knowledge persists and changes in species communities, and importantly, the role that theory of mind has to play in how that information can be delivered, and the impact of the fidelity of the information uptake. In this thesis, the social learning methods of most interest are emulation, imitation and teaching, as social learning strategies that differentially utilise theory of mind. Therefore their interpretations in the material record could serve as indications of those cognitive attributes.

In recent decades, and mostly in the last few years, experiments have explored how cultural transmission might have worked in human evolution by exploring the cultural transmission of lithic technology. The topic of language has been a feature of three of these experiments (Ohnuma et al. 1997, Putt et al. 2014, Morgan et al. 2015), though none have looked at the linking cognitive component of theory of mind to empirically tie the behavioural simulations to the cognitive implications of the material. Great progress is being made with these studies, and it is hoped that the experiment outlined in Chapter 4 will contribute further to the understanding of how experimental archaeology can elucidate the cultural transmission and cognitive behaviour involved in Palaeolithic technologies. The necessity for these more empirical approaches are necessary for progress in the discipline of cognitive archaeology, which has classically been dominated by conceptual argumentation and interpretation of material without empirical linkages, which leaves little in the way of building towards a fuller understanding. As discussed in Section 1.4, the dominance of the symbolic method in the literature of the last decades has meant that the Upper Palaeolithic is stressed as the source of modern human behaviour; more recently, with symbolic artefacts uncovered from the Middle Stone Age in Africa, cognitive complexity in human evolution has been pushed back. However, interpretations are limited to these late time frames, regardless of the complexities of human minds if they did not leave clear material traces for present archaeologists to interpret. The unfortunate issues with the symbolic method leave alternative approaches to Palaeolithic cognitive as highly desired, if only to corroborate current methods' interpretations. Theory of mind as a proxy, through the interpretation of social learning methods in the archaeological record, is argued to be a superior method of interpreting cognition in the record, and avoids many of the pitfalls of the symbolic method.

This thesis started with a research question and a chain of inference: if language ability and theory of mind ability have a predictive relationship (which was established in Chapter 2), and theory of mind enables certain modes of social learning, and social learning impacts copy error in material culture (which both of these points have now been established in the present chapter), is it possible to measure lithic assemblage variability to deduce social learning method, theory of mind ability, and therefore language ability amongst Palaeolithic stone tool makers? The next chapter will turn to the experimental design that sets out to answer the last part of this question.

Ohnuma et al. 1997 **Raw material:** Silicious shale **Number of Participants:** 20 **Learning Period/Training:** 8 hours **Method:** Teacher instructed **Group Conditions:** Silent, verbal **Replicative Goal:** Prepare core, detach Levallois flake **Dataset Size:** n/a (time to acquisition analysed)
Putt et al. 2014 **Raw material:** Flint **Number of Participants:** 24 **Learning Period/Training:** 5 hours **Method:** Teacher instructed **Group Conditions:** Silent, verbal **Replicative Goal:** Copy handaxe shape **Dataset Size:** 334 tools, plus flakes
Morgan et al. 2015 **Raw material:** Flint **Number of Participants:** 184 **Learning Period/Training:** 5 minutes **Method:** Iterative chains **Group Conditions:** Reverse engineering, Imitation/Emulation, Basic Teaching, Gestural Teaching, Verbal Teaching **Replicative Goal:** Detach flake **Dataset Size:** 6000 flakes
Schillinger et al. 2015 **Raw material:** Foam **Number of Participants:** 60 **Learning Period/Training:** 20 minutes **Method:** Either video instruction or none **Group Conditions:** Emulation, Imitation **Replicative Goal:** Copy handaxe shape **Dataset Size:** 60 tools
The study presented in this thesis **Raw material:** Porcelain **Number of Participants:** 20 **Learning Period/Training:** 5 hours **Method:** Either no instruction, video instruction or teacher instructed **Group Conditions:** Emulation, Imitation, Silent, Verbal **Replicative Goal:** Copy handaxe shape **Dataset Size:** 104 tools, plus flakes A summary of unique attributes of this study: • standardised porcelain blanks allow for controlled analysis of both tools and debitage. • only study that incorporates shape replication with separate emulation, imitation and taught methods • Explicit connections to cognitive mechanisms underlying methods

Table 3.1 A summary of attributes in previous lithic cultural transmission studies discussed in this chapter, as well as details of the study outlined for this thesis in Chapter 4.

Chapter Four.

Experimental Design and Methodological Framework

4.1 Introduction

This chapter details the structure of the experiment undertaken for this thesis in addition to the framework adopted for analysing the data collected. It will begin by describing the experimental framework, including information about participant make-up and selection, group categories, and the materials used throughout the experiment. This chapter then discusses expectations and the hypothesis the experiment is designed to test, before detailing the data collection procedure. This later section includes the variables considered for measurement and the recording methodology.

4.1.1 Ethical Considerations

Ethical approval for the study was granted by Ethics and Research Governance Online (ERGO), University of Southampton (Study ID: 18589). A risk assessment was submitted, because of the interaction with human participants, and approved. As the production of stone tools comes with certain health and safety risks, such as inhaling silica dust, or bruising and cuts from the tools, safety was a primary concern. It is noted that other experimental studies (e.g. Schillinger et al. 2015) avoided knapping stone because of the concern for health and safety. In an open and well ventilated area health risks are reduced to cuts and bruises and eye injuries. Proper safety equipment and preliminary training can overcome this. Participants were given an information sheet outlining the study with details on who to contact in the case of any problems arising during or after the study (see Appendix 1). A consent form which informed participants that they could discontinue the study at any time, for any reason, was also provided (see Appendix 2). Health and safety measures were outlined in the participation outline sheet, and repeated on the first day of the experiment. Gloves, goggles, and a leather chamois (to cover the thigh) were provided to lessen the risk of injury to damage to clothing. Participants were also requested to come wearing jeans if possible or other durable trousers, and to wear closed-toe shoes, for safety reasons.

4.2 Experimental Framework

The experiment was conducted in February, 2016, in Oxford, United Kingdom. It was an expansion of a smaller initial pilot study which took place in April, 2015, which had a smaller participant size and shorter participant training time, but was undertaken to ensure that the structure of the study would produce the desired data for testing the hypothesis within this thesis. The February 2016 study took place at Fusion Arts Centre over two weekends, where twenty participants learned to make handaxes on porcelain blanks, in one of four simulated social learning conditions: emulation, imitation, silent teaching, and verbal teaching. All groups replicated handaxes produced by the author. The experiment was designed so that all four groups would create materials that showed the effects that the specific simulated social learning method (the controlled variable) would have on the morphology of the handaxes produced.

The hypothesis to be tested was if modes of social learning used in the replication of lithic technology would differ by their range of relative morphological variability, due to varying levels of copy error from the replication process (Section 1.3).

In this case, the model to be replicated was a number (n = 5) of porcelain handaxes that the author had made. Aside from the emulation group, who replicated the model handaxes with no instruction, the other three groups were given access to knowledge of the manufacture of the handaxes, based on their particular simulated social learning environment.

4.2.1 Participants

In recruitment, participants were recruited by email, social media (Twitter and Facebook) and posters (Figure 4.1) advertising a knapping study, where people could learn to make stone tools in 'possibly unconventional situations'. Eligible participants were to have never knapped before, to reduce the chance of varying skill biasing results. Participants were also encouraged not to research knapping in any form during this study, to also reduce bias. While studies have used financial incentives as a form of remuneration for participating in the experiment (e.g. Morgan et al. 2015), it was deemed more desirable to offer a starter knapping kit to the sum of £5 (Figure 4.2). This kit included: 1) gloves, 2) goggles, 3) a leather chamois, 4) a quartzite hammerstone, and 5) a flint nodule (sourced from Sombourne Chalk Quarry, Hampshire). A sheet providing more information on how to pursue knapping as a hobby (local clubs, recommended books and flint sources) was provided. From recruitment, fourteen women and six men were accepted for the study. The majority of participants were undergraduate students or Further Education students from the University of Oxford or Oxford Brookes University. The participants were of varying age, body size and handedness. These variables were not controlled for a number of reasons including ethical considerations and experiment feasibility.

Before starting the experiment, participants were given a short introductory lecture from the author on stone tools and in particular handaxes, and were able to handle a fine modern flint handaxe made by expert knapper John Lord. This introduction was designed to equalise their beginning knowledge state and exposure to stone tools as some might have had some exposure to information about Palaeolithic tools through the media or other outlets.

Figure 4.1. An advertisement circulated on social media for the recruitment of participants in the thesis experiment.

Figure 4.2. Example of a knapping kit participants were given as a thank-you for taking part in the experiment.

Participants were assigned to one of four groups, based on participant availability. Groups were not randomly assigned however participants did not know what group they were signing up for; they only knew what timeslot they would be joining, and that they would be learning to make stone tools from one of four different and possibly 'unconventional' ways. No requests were made to be a specific learning style group.

4.2.2 Structuring the Experiment

The four groups in the experiment each differed in the information given to them about how to make a handaxe, depending on which group they were in. After receiving their group specific instruction each participant was given up to fifteen minutes to replicate the model and produce a handaxe. Fifteen minutes was deemed sufficient time to complete a handaxe. This procedure was reproduced seven times (each time being referred to as an 'attempt'), as this exhausted financial and time resources.

The requirements to make a successful handaxe in this study were simplified to a minimum of tasks. Instead of blocks of flint, whose size and shape could not be controlled for, participants knapped porcelain blanks especially made for the experiment. These are discussed below (Section 4.2.3). Knapping of the porcelain blank required fewer skills than the creation of a 'conventional' handaxe, and excluded much of the *chaîne opératoire* that past populations undertook. As detailed below, the porcelain blanks were easier to fracture than flint as less force is required. Secondly, the blank shape already features an acute angle around the perimeter, allowing a more desirable geometry for flaking. The blank retains no 'cortical' surfaces or shape irregularities. Raw material selection was also not a factor, nor was quartering a suitable nodule to create an appropriate blank (which can then be shaped into a handaxe).

Participants were to replicate a successful handaxe simply by creating a bifacial edge around the blank perimeter, and shaping the tip into a point like the model handaxes. To replicate the model participants needed to flake the perimeter by first striking the flat face of the blank, detaching parallel invasive flakes approximately 10mm in from the edge. After this, the blank could be flipped over and the perimeter could be knapped using the platforms made from the negative scars of the previously produced flakes. The same effect could also be made by alternately flaking either side of the handaxe, flipping the handaxe after each flake, using the previous negative scar as a platform for the next removal. Both methods result in a sinuous edge running around the perimeter of the piece (but all model handaxes and demonstrations were created with a parallel flaking method). Edges were then evened out so a straight edge profile was achieved, and a tip formed at one end of the piece by removing more material on both faces. A pointed handaxe shape was chosen as the model shape for ease of orientation. This approach allowed for greater control of the starting variables common to all knapping episodes. Reducing this 'noise' focused analysis on differences that would reveal similarities and differences between the knapping in each group. See Appendix 3 for samples of handaxes produced in each group.

Details about the specific simulated social learning environment are provided for each group below.

4.2.2.1 Group 1 (Emulation)

In this group participants learned to make a handaxe by replicating the end products of the manufacture sequence; that is, the model handaxes made by the instructor (Figure 4.3). This group had to replicate a handaxe without knowledge of the manufacturing technique, using clues only from looking at the finished model/tool and its associated debitage, though they knew that it had been created by striking the blank with the small hammerstone they were supplied with.

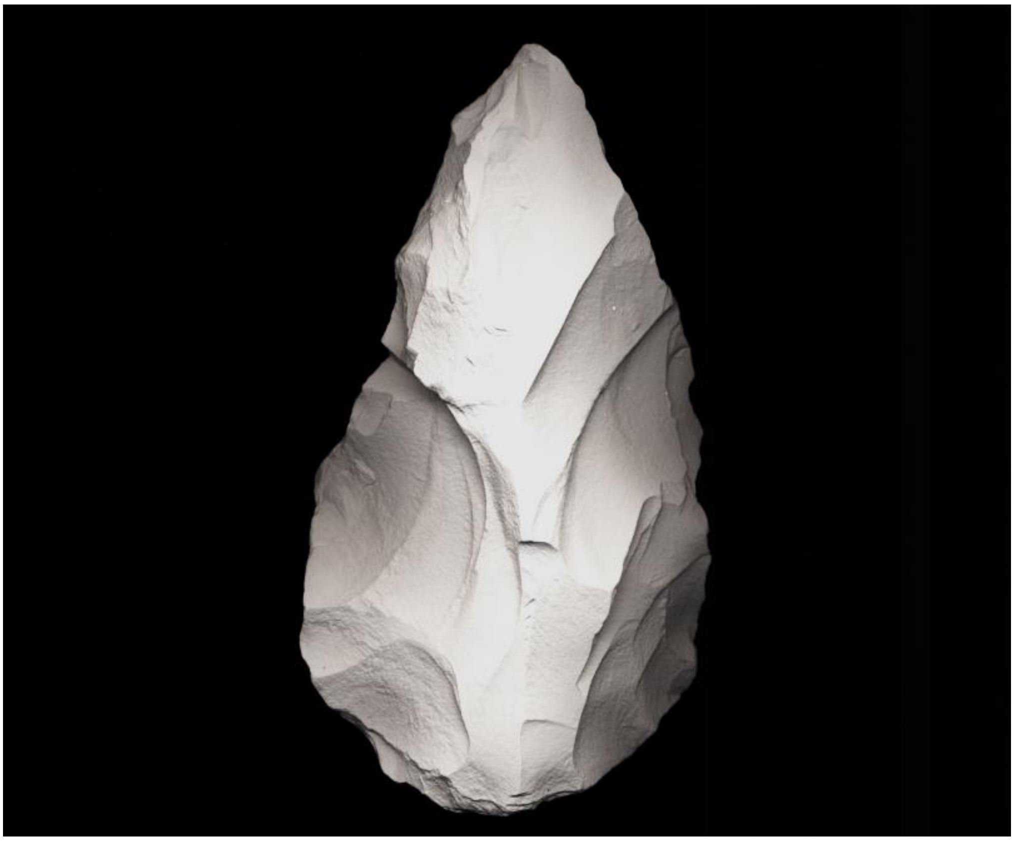


Figure 4.3. One of the five sample handaxes made by the author and used as a model in the experiment for the participants to replicate (Sample 4).

The protocol for Group 1 (Emulation) was as follows. Participants in this group were sat facing away from each other, towards the wall, and they were instructed not to watch and copy other participants (however discussion between participants about any other subjects was allowed), to minimise the risk of imitating what others were doing, and so to artificially construct an emulative situation. Participants were then given a sample handaxe and its debitage, and were instructed to inspect the contents before attempting to recreate the sample themselves. The sample was kept with them for reference. Often the participants tried to fit the pieces back together in an effort to understand the manufacture sequence better, and reverse-engineer the process. As this group had only the sample, as well as their own trial-and-error, this group paid more attention to the sample assemblages than did the other groups, and the other groups did not try to refit the pieces as extensively as this group did. Each participant was then provided with a porcelain blank and had fifteen minutes to try and replicate the model. After the fifteen minutes, the handaxe and associated debitage were put into a labelled bag for later analysis. While the protocol ensured an environment where inter-participant copying was controlled for, participants could hear other knappers' percussive motions, which may have provided hints as to what method others were trying. This includes rapid percussions and grinding/abrasion.

After the experiment was completed, the group was shown videos of the creation of the sample handaxes they had been copying, as a debrief.

This protocol created an environment of 'end-state' copying, known in the relevant literature as emulation (Huang and Charman, 2005). As discussed in Chapter 3, this group would not be using theory of mind to aid in the handaxe manufacture. In an emulative situation, an individual would not learn from watching the processes of others. The process others use may be visible, but are not encoded or connected with the goal. In order to follow the instructions, Group 1 (Emulation) participants were all capable of advanced theory of mind, and the communication skills necessary to understand the instructions and language used within the experiment. This is why the group is a 'simulation' of emulative learning: as the tool-making process was not accessible to participants, it was therefore impossible for the participant to knowingly replicate another's process to make the goal. Nowhere in their knapping were they using information inferred from viewing and replicating the process conducted by someone else (except for what they could hear, noted above). This protocol is in line with Morgan et al. (2015) and Schillinger et al. (2015) in their simulated emulative learning environments.

4.2.2.2 Group 2 (Imitation)

In this group participants learned to make a handaxe by replicating the process seen through watching a short video of the author making the model handaxes used in the experiment (Figure 4.4).

The protocol for Group 2 (Imitation) was as follows. Participants were sat in a semi-circle, around a laptop which played videos and sound (7-8 minutes in duration) of the handaxe manufacture process, and then given a handaxe blank to replicate the process. The videos shown include the knapping of a model handaxe, from start to finish, filmed at various angles. During the videos, the instructor did not gesture towards the camera or try to slow or emphasise movements in any way. The video-based lessons showed the knapping of the sample handaxes to simulate participants 'spying' on the model, where learners could attempt to learn from someone unaware that they are transmitting knowledge. Participants were instructed they could make handaxes by watching the video, as well as watching other participants within the same group. Participants were requested not to discuss their knapping activities or to help each other (to prevent teaching or the transmission of cultural knowledge transmitted by social learning requiring third-level intentionality). This group also had fifteen minutes to produce a handaxe after watching the short video. While the handaxes were being made, videos were left on a loop in the background for reference. As in Group 1 (Emulation), the model handaxes and debitage were available for participants to examine throughout the process.

Figure 4.4. A photo still from one of the videos shown to participants in Group 2 (Imitation), Group 3 (Silent) and Group 4 (Verbal).

This protocol created an imitative environment, as opposed to a teaching environment as while the process of the manufacturing sequence is being exhibited in the videos, the demonstrator was not intentionally conferring knowledge. The use of a video instead of a non-interactive live demonstration was to control any involuntary emphasis a teacher may have placed on their actions, for example reacting to participants struggling, or needing emphasis or explanation of a particular aspect of the knapping. Clever Hans is an example often used in psychology research showing how unintended cues can be given that may bias an experimental setting (Pfungst, 1911). This is why in experiments with children and infants where the participant is held by the parent or sat on their knee, the parent is often blindfolded so they cannot offer any unintended cues during the experimental process. Schillinger et al. (2015) also used video in their imitative learning situation, possibly for these reasons.

This protocol was designed to only allow second-level intentionality (basic theory of mind) to aid in the manufacturing of the handaxes (see Chapter 3). In this, the participants could theorise about how the process was connected to production of the handaxe in the video. The participants, because they were all able to use theory of mind in complex ways, had no problem watching the video and theorising which intended action created which intended consequence, because all participants were typical adults who had functional theory of mind abilities (this was obvious and did not need to be tested for). For example, in one of the videos the instructor brushes flakes off their chair; the participants knew that this was not an integral step in the completion of the handaxe, and was not part of the process. However, the striking of the porcelain blank with the hammerstone was seen as significant to the task.

4.2.2.3 Group 3 (Silent)

In this group participants learned to make a handaxe by replicating the process seen in a live demonstration by the author. However, no verbal language was used in reference to the task, and so it was referred to as the 'silent' group.

The protocol for Group 3 (Silent) was as follows. Participants were sat in a semi-circle, facing the instructor who was also seated. The participants witnessed a demonstration of the handaxe manufacturing technique, where the instructor drew attention and pointed to key aspects of the technique. The author/instructor slowed and emphasised certain aspects of the procedure, for their benefit, before each of the seven attempts. Participants and the instructor interacted through gesture, correction, and emphasis, without any verbal communication (and none of the participants communicated using sign language). After this, participants were given fifteen minutes to replicate the model. They also had access to the model handaxes used for reference in other groups, and the videos used in Group 2 (Imitation) were also played on loop for reference. Group 3 (Silent) were also instructed that they could copy the methods they thought others were finding successful, just as Group 2 (Imitation) were instructed. This group was intended to be similar to the verbally taught group, with only the verbal element missing. It was not non-linguistic, as some gestures were culturally learned and intentionally communicative symbols (such as a ‘thumbs up’), and like Group 4 (Verbal), were utilising a social learning method which requires third-level intentionality.

Group 3 (Silent) was included in this experiment in order to replicate studies such as Ohnuma et al. (1997) and Putt et al. (2014), in which artefacts were compared between verbally and non-verbally taught groups. Both studies concluded that because there was little difference between the two groups taught with silent and verbal teaching that language may not have been a requirement for complex knapping skills in the Palaeolithic. However, silently teaching a group is not controlling for language. Participants would still be using communicative and cognitive skills only possible with their developmental context of being complex language-using individuals. Both Ohnuma et al. (1997)'s and Putt et al. (2014)'s studies allowed gestural communication and other communication methods that required third-level theory of mind, as well as assumption of other's linguistic competence, which Chapters 2 and 3 laid out as requiring a high level of linguistic ability to have developed in the first place; a hypothetical community that could teach, could, under no circumstances, also be non-linguistic (except perhaps in the case of post-developmental impairment). In this experiment, both Group 3 (Silent) and Group 4 (Verbal) were using third-level intentionality (and, if necessary, higher theory of mind, although the opportunity most likely did not arise, or aid in the information transfer of handaxe manufacture). Ohnuma et al. (1997)'s and Putt et al. (2014)'s studies then simply controlled for how much information was being delivered verbally, which, while interesting, and partly why is it also replicated in this study, is not appropriate for inferences about hominin language necessity in Palaeolithic stone tool cultural transmission.

4.2.2.3 Group 4 (Verbal)

In this group participants were taught similarly to Group 3 (Silent), but with the addition of verbal communication (Figure 4.5).

The protocol was as follows. Participants were sat in a semi-circle, facing the instructor who was also seated. Before each attempt, as in Group 3 (Silent), participants witnessed a demonstration of the handaxe manufacturing technique, where the instructor drew attention and pointed to key aspects of the technique, in addition to describing the techniques verbally. Like the other three groups, this group was given an opportunity to inspect a sample handaxe and its associated debitage, followed by videos of the knapping of the sample handaxes, in addition to a live demonstration as with Group 3 (Silent). Fifteen minutes were allocated to replicating a handaxe.

Figure 4.5. A photo of Group 4 (Verbal) which shows their semi-circular seating arrangement and debitage collecting at their feet.

Group 4 (Verbal) represents a teaching-learning situation that utilises third-level theory of mind, because the instructor is intentionally conferring knowledge to the participants who are able to encode the information conveyed to them about the process of handaxe manufacture. As in Group 3 (Silent), participants were not limited to using more complex levels of theory of mind, however those uses most likely did not arise, and if they did they would not aid in the production of handaxes for the experiment.

4.2.3 Experimental Materials: Further Information

To ensure a standardised practice throughout the experiments a number of factors were controlled and standardised to the highest extent possible. Throughout all four groups, small rounded hammerstones (c. 150-200g) were used. They were somewhat standardised in shape and of the same material (quartzite), all sourced from the same quarry in Oxfordshire. The experiments were conducted in a well ventilated large room, so that weather, heat and lighting were consistent and not an issue for the participants knapping. Ground sheets were placed underneath the participants so that all material (from flakes to dust) could be collected and catalogued appropriately.

One of the main controls used throughout was with the adoption of porcelain as a knapping medium (see below). The porcelain blanks in this experiment were made of standard porcelain purchased from Bath Potters' Supplies (called 'Standard P2 Porcelain'), and was fired at 1250°C. All blanks were produced from the same single mould, which was made from Hobbycraft air-dry clay into the negative shape of an idealised flake for handaxe manufacture (see Figure 4.6), flat on the bottom, and domed on the top, in an ovoid shape. 1000g of wet porcelain clay was then weighed and pressed into the mould to shape it, before being lifted out and set aside for drying, which assured a standard shape and size. After firing, due to water lost during the heating process, the porcelain pieces each weighed 880g.

All model handaxes (n = 5) produced by the author were of the same material as used in the experiment (porcelain blanks), and reduced using the same hammerstones present. The model handaxes were of similar size and shape, corresponding to a pointed-type (Figure 4.3). The production of each of the model handaxes were recorded and used throughout the experiment in Group 2 (Imitation), Group 3 (Silent) and Group 4 (Verbal).

Figure 4.6. Air-dry clay was formed by hand into the negative shape of an ideal handaxe blank, and was used to reproduce all of the porcelain blanks and model blanks for replication in the study.

4.2.4 Justification of Handaxes as the Artefact Target

The exercise could have been carried out with participants trying to make any number of media (such as drawings or sculpture), as long as they offered participants enough variables where they could vary in terms of copy error in reproducing the model pieces, so that the comparison in ranges of variability could be measured to the test the hypothesis (see below). Knapping a handaxe was chosen because lithic technology is the goal application, and there are few other simple, cost-effective reductive technologies that participants would have little information or experience with, and yet could be learned easily. It is important to note however that this is not an experiment to test how experimenters learn how to knap a handaxe; much of the skill involved in handaxe making and bifacial knapping has been purposefully omitted from this experiment in order to target the transmission of select variables. Another reason for choosing a handaxe shape is that Schillinger et al. (2014) notes how additive versus reductive technologies may differ in their morphological tendencies, and Kempe et al. (2012) suggests that reductive versus additive technologies may have different cultural mutation rates.

Choosing knapping, and specifically the knapping of handaxes, also allows discussion of this study alongside numerous other studies of handaxe variability, for reasons which Lycett (2015) has called handaxes a 'model organism', because of their utility and common use as case studies regarding archaeological materials. As this thesis considers the origins of language, handaxes are a suitable technology that would have likely been used during language's early evolution considering its multi-species and wide temporal usage (Lepre et al. 2011; Lycett and Gowlett, 2008). The application of this methodology wishes to be applied to interpretation of archaeological lithic assemblages, and so using a lithic technology removes confounding variables that can be introduced using other materials. The experiment conducted here therefore uses the handaxe form to study the effects of social learning on morphological variability.

4.2.5 Justification of Porcelain as a Knapping Medium

The raw material used in this experiment was porcelain, which vitrifies when fired and features similar fracturing and mechanical properties to high-quality flint (Khreisheh et al. 2013), although in truth it is even easier to detach flakes from (requiring less force). Porcelain features conchoidal fracturing, just as flint, obsidian, or other silica-rich knapping materials do; it can be used with hard-hammer or soft-hammer percussion, and even pressure-flaked effectively (see images in Khreisheh et al. 2013). Porcelain, as opposed to flint, will not contain frost-fracturing, fossils, geodes or other structures which hinder the knapping process. Furthermore, another benefit is that porcelain clay is relatively inexpensive and has no waste debris in the manufacturing of the blanks. Porcelain was also chosen because it can be moulded into identically sized blanks before being fired, which introduces a further experimental control and allows for the comparative morphological measurements of the tools and products of the knapping sequence.

Porcelain can also be seen as safer to use than flint since it does not require as much force to detach a flake. Porcelain can be held more comfortably on a participant’s lap, without causing any bruising to the thighs or the need for padding, and it meant that any accidental slippage and hitting of oneself would hopefully not be hard enough to cause serious injury. Anecdotal evidence and experience with the material during the experiment also suggests the edges produced are not as sharp, and are therefore less likely to cause injury.

Figure 4.7. Spear point made from a porcelain insulator from northern Australia, with adhesive still adhering at the base. Accession number: 1900.55.42. Pitt Rivers Museum.

Porcelain has been used in certain Australian communities as a natural stone substitute for making tools (Figure 4.7), and by some knappers in knapping practise (Whittaker 1994). This led archaeologists at the University of Exeter to develop the use of porcelain as a raw material in experimental archaeological studies during the Learning to be Human project (Kreisheh et al. 2013). Otherwise, its use in experimental studies has been limited. It was used as a raw material in a recent PhD thesis on skill transmission of blades and handaxes (Page 2014), a recent Bachelor's thesis looking at snapped Clovis blades (Klemenic 2012, Kreisheh et al. 2013), and in a recent Bachelor's thesis looking at edge angles of Clovis adzes (O'Leary 2012). Further research into the qualities of porcelain and how it compares to natural stone might encourage its future use as a controllable raw material in experimental archaeological studies involving stone tools.

One of the key problems with the experimental design of other experimental knapping studies concerns running out of the raw material used. For example, Ohnuma et al. (1997) ran out of the initial raw material that they were sourcing locally from a river bank, and needed to move to another area in order to conduct the non-verbal group's part of the study. This meant that the two groups had used different raw materials that could have affected the outcomes. Similarly, in Putt et al. (2014), the flint that was being used ran out and another type was brought in during the study. The use of porcelain blanks in this study allowed the confident knowledge of the consistency of material, and the replicability of it in the future as well. In Ohnuma et al. (1997), Putt et al. (2014), and Morgan et al. (2015), participants used starting-blanks which were not standardised, which affected the size and shape such that shape analysis could not be effectively conducted. Schillinger et al. (2015) overcame this by using garden foam bricks to ensure that all participants were using materials that were standardised, and they were able to conduct an effective shape analysis. However, the nature of the garden foam meant that the 'debitage' was unable to be included in the analysis. Porcelain, then, avoids a number of the problems faced by these previous studies and is recommended here for future experimental knapping studies.

4.3 Expectations and the Null Hypothesis (H_0)

The null hypothesis was that experimental handaxe assemblages will feature no difference in their range of morphological variability (i.e. outline shape) when different simulated social learning methods are used as a controlled variable.

It was anticipated that as the different simulated social learning methods would each convey different resolutions of information about the task that the participants would be limited in the amount of knowledge that they could encode, replicate faithfully, and thus differing levels of convergence on the target form.

Many of the expectations derive from observational and anecdotal evidence, and experimental protocol, coupled with knowledge gained from previous studies e.g. Schillinger et al. (2015) and Morgan et al. (2015).

For instance, in Group 1 (Emulation), after the participants had completed the experiment they were shown the videos of the handaxes being manufactured as a debrief, and their reactions were of genuine surprise. The participants from Group 1 (Emulation) had reported confusion as to the process they were trying to replicate and they remained unsure that they had worked out the best way to copy the piece. This strengthened the expectation that emulative groups would exhibit the largest morphometric ranges due to the varying techniques employed. Group 1 (Emulation) was also the only group that did not have a demonstrated method to sit, hold the blank, and strike with the hammerstone, and their methods of holding the porcelain blank was varied. The varied physical stance in Group 1 (Emulation) was expected to influence variability immediately as other groups had already converged in their physical stance, and hence range of motion and physical behaviour with the blank. Conversely, the videos (and live demonstrations) offered the participants a model of how to sit and hold the blank, how to strike with the hammer, and with what approximate force, the arc of the hammerstone, and other clues they could pick up visually. These cues were not visible to Group 1 (Emulation), who had to decide what was best for themselves through trial and error.

In the same vein, Group 3 (Silent) and Group 4 (Verbal) were exposed to more knowledge about to the task that remained opaque to Group 2 (Imitation), even though the imitative group were able to observe the knapping sequence through the video. The two taught groups were shown exactly how far in from the edge to strike the porcelain blank with the hammerstone in order to detach a desirable flake. While Group 2 (Imitation) could approximate this through the video, it was stressed to the taught groups the difference between a flake with a large platform, and a flake with a small platform, and the difference this caused to the size of the detached flake, allowing greater predictive control.

Group 4 (Verbal) were the only group where it was simple to warn participants not to do certain actions detrimental to the knapping process. In Group 3 (Silent), without the aid of language, it was near impossible to deliver foresight about what not to do in the knapping process e.g. failing to support the blank during flake detachment, so that the blank does not split. It was therefore expected that Group 4 (Verbal) would have the smallest handaxe breakage rate.

It was also expected that the quality of knowledge conveyed about the knapping procedure would affect the rate of learning to detach successful flakes in a predictable manner, which would then affect the overall morphology of the handaxes. Group 1 (Emulation) received the least amount of information on flaking technique and was so expected to be poorest, while Group 2 (Imitation) would obtain a number of visual cues on the knapping technique, which allowed participants in this group to learn faster and become more efficient in flaking ability. The taught groups received the most specific information on flaking technique, including information on platform size, and thus would learn flake detachment faster, and finish the experiment at a higher level.

4.4 Data Collection Procedure

To assess variability among the four groups a number of variables were selected and recorded from the handaxes and their associated debitage through a statistical framework. This selection will detail why each variable was chosen, and the method undertaken for recording those variables.

4.4.1. Selecting Variables for Measurement

In their entirety, three categories of variables were considered for assessment: 1) knapping intensity and reduction, 2) attributes pertaining to flake artefacts, and 3) final handaxe shape (as determined through a geometric morphometric framework).

To assess variation among the four different simulated social learning environments, aspects of knapping intensity and reduction were selected for investigation, given the expectations outlined in Section 4.3 that interaction with the knapping medium would vary given their differing consistency in the production of flakes. The main manifestation of this would result in differing amount of material being detached, and therefore it was deemed appropriate to investigate the amount of material reduced, and the degree of knapping undertaken. This was assessed through two variables: weight lost and scar count. Weight lost would be calculated as the weight of the original blank (880g) minus the weight of the resulting tool. Scar count would be calculated as a count of all negative scars measuring greater than 2cm in axial length. 2cm was deemed suitable, given its use elsewhere as a measure of counting flakes in Palaeolithic contexts (e.g. Roberts and Parfitt 1999, p. 316).

Another method of assessing variation among the four different simulated social learning environments was to analyse the material detached from each knapping event. This would directly link to the expectations of knapping skill (Rein et al. 2013), that emulative learning would produce both poor and varied knappers (who would produce smaller and fewer flakes in number), and the taught groups would be the most successful flakers (who would produce the most flakes). Two measures were recorded: flake count and flake weight. Broken and unbroken flakes over 2cm were included in this analysis.

A final method of assessing variation, irrespective of size, was to consider the shape of the handaxes produced through a geometric morphometric framework. Here, shape is defined as the total amount of information that does not vary after translation, rotation and/or size (Small, 1996). Shape was determined as important to investigate because the myriad of variables and physical determinants affecting the blanks through the knapping sequence were expected to influence the morphology of the end tool. In this, shape is a measure which captures the most information about a handaxe, when technological analyses are excluded.

To capture 'global shape', i.e. the overall topography of the handaxe, three perspectives were utilised: the 1) planform (planar), 2) right lateral (side), and 3) superior (top-down) views. This was to ensure that the most information about shape possible could be analysed. A full three-dimensional approach was not feasible for a number of considerations, including the lack of appropriate equipment (a 3D scanner), the difficulty in 3d scanning white porcelain, which exhibits a reflective surface, and the difficulty in 3D landmarking crude handaxe shapes without a continuous edge..

4.4.2. Recording Methodology

To measure flakes and flake scars, digital callipers were used. These had a resolution of 0.01mm, and an accuracy reading of ± 0.02mm for measurements < 100mm, and ± 0.03mm for measurements >100-200mm. In measuring the weight of the handaxe produced, and the weight of the flakes greater than 2cm digital scales to 0.01g resolution were also used.

To record shape, each handaxe was first scanned at 300dpi with a Plustek Opticbook 3800 Flatbed Scanner, and processed in CorelDraw X7. To ensure that handaxes were positioned correctly on the scanner, fixed weights were used to support the handaxe during the scanning procedure. For each handaxe, the planform, right lateral and superior views were digitised. Images were then traced using CorelDraw X7's 'Trace Outline' function, with the original image deleted (see Figure 4.8 for an illustration of the outline and landmarking procedure). This was in order to reduce pixel noise. The outline was then screened to ensure that all curves around the shape were complete; incomplete curves were closed using CorelDraw X7's 'Close Curve' tool. Following this, the thickness of the outline produced was set to a thickness of 1 pixel to again reduce pixel noise.

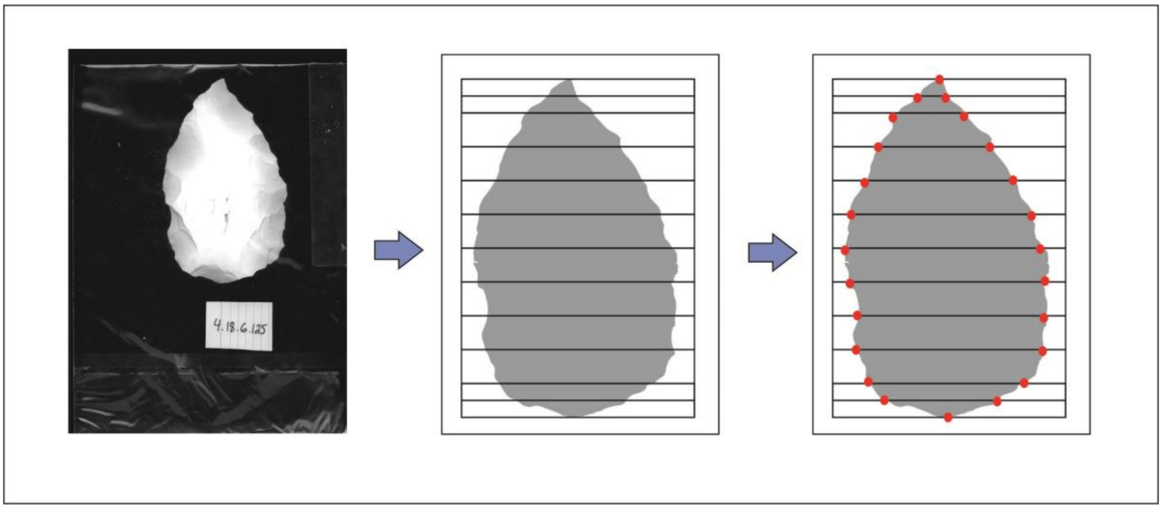


Figure 4.8. A depiction of the procedure to create outlines of the handaxes images and apply the grids for landmark placement.

For any analysis incorporating a geometric morphometric framework, standardisation in handaxe orientation is essential so that comparisons in shape throughout an assemblage are homologous (McPherron and Dibble, 1999; Lycett et al., 2006; Costa, 2010). For the planform and lateral views, handaxe outlines were oriented along their axis of symmetry, a method of orientation first exemplified by Callow (1976) and elaborated by McPherron and Dibble (1999) and Costa (2010). For the superior perspective, a longitudinal perpendicular view of the planform was obtained.

To examine shape, semilandmarks were obtained and defined by the shape's geometry. In CorelDraw X7, a 10 x 2 grid (1 pixel thick for consistency) was placed onto the maximum dimensions of the planform and lateral views, with a 2 x 10 grid placed onto the maximum dimensions of the superior view. This allowed the placement of twenty semilandmarks, longitudinally equidistant, positioned at every 10% of the handaxe's width (with an analogous scheme for the superior perspective). A further 1 x 2 grid was superimposed on the handaxe tip and base, for both planform and lateral views, in order to better characterise changes in the shape's extremities. A similar methodology was also adopted for the superior view on its extremities. The handaxe outline and the superimposed equidistant grids were then exported as a single .jpeg file, ready for landmark placement.

All .jpeg files were collated into three thin-plate spline (TPS) data files in tpsUtil v.1.69 (Rohlf, 2016a), one for each of the perspectives considered. Two-dimensional (x, y) semilandmarks were then plotted where the handaxe and grid intersect, using the 'Digitize Landmarks' function in tpsDig2 v.2.27 (Rohlf, 2016b). Each handaxe therefore features seventy-two semilandmarks, twenty-four on each perspective. All artefact identification numbers were appropriately relabelled, with the teaching method used, via Notepad.

All data was transcribed into PAST v. 3.12 (Hammer et al. 2001), for subsequent statistical analyses.

4.4.3 Analytical Procedure

Summary statistics for each variable (excluding shape) were first recorded. This includes the minimum and maximum values, the mean and median values, the standard deviation and the coefficient of variation (CV), that is, the ratio of the standard deviation to the mean. These summary statistics were recorded for each variable and are presented in order to give a broad characterisation of the values before discussing the variable in greater detail. Descriptors of central tendency and dispersion were also recorded through the creation of a box-plot, to better visualise differences in the range of morphological variation for the four simulated social learning groups. This is important as the range of morphological variation is the key variable this thesis aims to explore, as per the hypothesis.

All continuous variables were first assessed for their normality through a Shapiro-Wilk test (calculated in PAST). This normality test was used because it determines what statistical measures are appropriate, as some statistical measures can only be used on values with a normal distribution for example (Shapiro and Wilk 1965). If normal, the groups in question were subject to a test for equal means (two-sample *t*-test), and a Mann-Whitney U test for equal medians if a normal distribution could not be assumed. These tests allow us to see how likely it is that a sample from two groups were from populations with the same distribution – i.e. it tells us how different the groups are (Fay and Proschan 2010). These tests were undertaken to test each of the simulated social learning groups against one another for statistical significance. Statistical significance is here recorded to a 95% confidence level (i.e. α: 0.05), the most commonly set significance level (Craparo 2007). Groups were then analysed, separated by group, to look at variability by attempt (i.e. handaxe attempt) number. Given the size of the dataset in question, visual descriptors through a box plot were the best form of assessing morphological variability, and statistical tests could not be reliably performed for the attempt number simply because at this resolution there were sometimes only three artefacts in a group.

Following the collection of semilandmark data, all x,y coordinates were transformed into Procrustes coordinates, following the iterative algorithm of Procrustes fitting (Rohlf and Slice, 1990). The process of Procrustes Superimposition removes size, translation and orientation, allowing an analysis of shape independent from all other variables. In effect, the x.y landmark outlines of each artefact have been placed on top of each other, resized to be the same size as each other, and oriented all in the same way – 'tranformed' so that the only element being looked at is shape (Figure 4.9). These Procrustes coordinates were then analysed through a Principal Component Analysis (PCA) in order to analyse the main sources of shape variation throughout the dataset (Hammer and Harper 2006). A PCA finds which hypothetical variables account for the most variance in the data; in this case, the variables are the relationships between the x,y coordinates. This is used for the analysis in this thesis because it allows for a visualisation of the range of these different shape variables. Percentage values of each source of variance (called a principal component) were retained in order to gauge a sufficient analysis of overall shape. As the first three sources of variation account for around 70-80% of accumulated shape variance, only these are outlined. The data was displayed as box plots, in contrast to bivariate axes as this visual descriptor is consistent with analyses in this thesis, analysing the degree of variation in morphological change among the four simulated social learning groups. Reconstructions of the main sources of shape variation are also displayed. To test for statistical significance in shape among the four groups, the sum of principal components totalling 95% of cumulative shape variance were analysed through a Multivariate Analysis of Variance (MANOVA), with significance again recorded to a 95% confidence level (a MANOVA test the means of several multivariate groups, Hammer and Harper 2006; it is used here as a way of telling if the groups' values were statistically different).

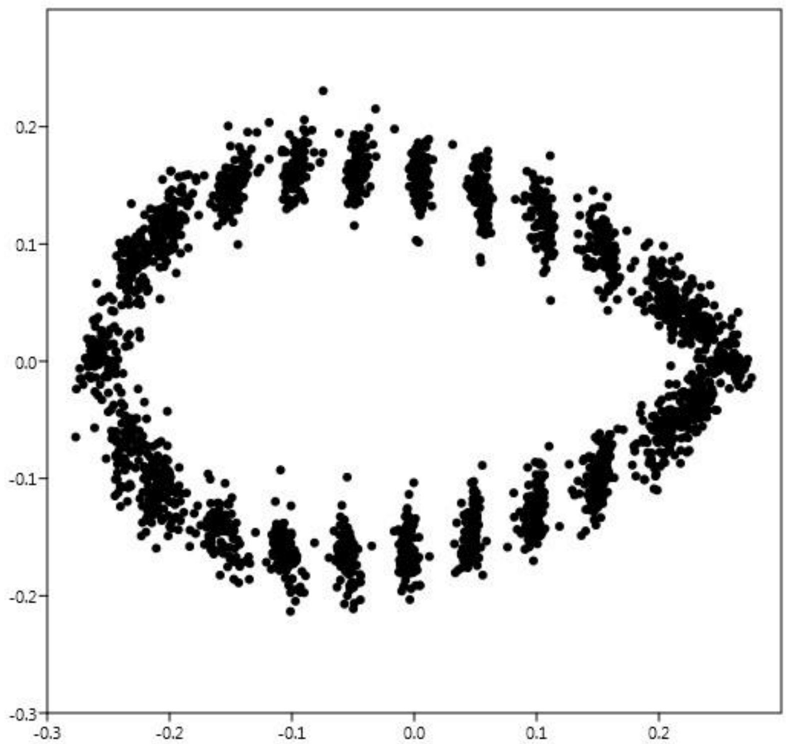


Figure 4.9. The x,y coordinates of the planform handaxe landmarks after Procrustes transformation.

4.5 Conclusion

This chapter has conveyed the experimental design, including ethical procedures undertaken and considered, the framework and structure of the experiment including characterising the participants and the four simulated social learning groups, and provided information on the experimental materials. It also discussed the justification for using handaxes as a target for replication, and the choice of porcelain as a knapping medium. Expectations of the outcomes of the experiments along with the statement of the null hypothesis were also outlined, which will be important for the discussion (Chapter 6). The procedure for data collection was also explained, including the justification for selecting specific variables for measurement, and the recording methodology that was then undertaken, before illustrating the analytical procedure used in this thesis. This has given an overview and justification of the structure of the experimental design so that the results and following discussion are given thorough context.

Chapter Five.

Results of the Knapping Experiment

5.1 Introduction

This chapter presents analyses from the experiment detailed in Chapter 4, examining individual aspects that were deemed likely to show morphological variability among the experimental handaxes and their associated debitage. These variables include traditional morphometrics examining knapping intensity and mass reduction (weight lost, scar count and scar density), and flaking attributes (flake count and flake weight). A geometric morphometric analysis uses Principle Component Analysis for visualisation of the range of shape variation.

5.2 Data Included in the Analysis

Throughout the experiment, ten instances failed to produce a handaxe, here meaning a porcelain blank which had either been worked bifacially, or a prominently large piece with shaping effort having been clearly applied by the participant. For example, sometimes the porcelain blank broke into many smaller chunks, resulting in few flakes, and bearing little resemblance to the intended replicated tool. These attempts were therefore not included in the analysis (see Table 5.1 for a breakdown).

In addition, on the second day of the experiment, one participant in Group 3 (Silent) failed to attend the experiment; their attempts were omitted from the analysis (to prevent this group having pieces weighted towards earlier, less practised manufacture from the first day). Hence, the data analysis totals 104 handaxes from nineteen participants. For a breakdown of numbers of artefacts by group and attempt, see Table 5.2.

	Total attempts made	BASM examples	Failed examples	Breakage Total
Group 1 (Emulation)	30	14	5	19 (63.3%)
Group 2 (Imitation)	30	6	4	10 (33.3%)
Group 3 (Silent)	24	5	0	5 (20.8%)
Group 4 (Verbal)	30	5	1	6 (20.0%)

Table 5.1. Overall breakage rate of handaxes. Broken and subsequently modified ('BASM') denotes pieces that broke and a tool was created out of one of the larger pieces. 'Failed' denotes when a participant broke their porcelain blank to the point there was not a recognisable 'tool.

	Attempt 2	Attempt 3	Attempt 4	Attempt 5	Attempt 6	Attempt 7	Total
Group 1 (Emulation)	4 (2)	5 (3)	3 (1)	4 (2)	4 (3)	5 (3)	25 (14)
Group 2 (Imitation)	5 (1)	4 (0)	4 (3)	4 (1)	5 (0)	4 (1)	26 (6)
Group 3 (Silent)	4 (1)	4 (0)	4 (3)	4 (1)	4 (0)	4 (0)	24 (5)
Group 4 (Verbal)	4 (0)	5 (2)	5 (2)	5 (1)	5 (1)	5 (0)	29 (6)
Total	17 (4)	18 (5)	16 (9)	17 (5)	18 (4)	18 (4)	104 (31)

Table 5.2. Total number of handaxes included in the data analysed by group and attempt number. Brackets indicate broken and subsequently modified (BASM) examples that were omitted in certain analyses (see below).

Because of the different numbers of handaxes included in each group, one criticism is that it could skew the data analysis. Group 1 (Emulation), for instance, had a total of five handaxes omitted due to failed attempts, while Group 4 (Verbal) only had one handaxe omitted. Group 3 (Silent) had no omissions that were due to failed attempts; but, with one omitted participant's data, the group remains the smallest overall. Group 1 (Emulation), Attempt 4, contains only three handaxes out of a potential five. These differences between groups are small, and are anticipated to not introduce much bias in the analysis. However, since the hypothesis being tested is that Group 1 (Emulation) will have the largest amount of variation and Group 4 (Verbal) will have the lowest, the bias (if any is introduced) will not result in a false positive (Type 1 Error), but will only introduce 'noise', as there is now potential for more variation in the higher fidelity groups which contain more handaxes. The higher quantity of data in these higher fidelity groups will only dampen any expected trends, and not accentuate them artificially.

Unintended breakage and subsequent modification of a chunk of the original blank happened more often in earlier attempts than in later attempts across all groups. This created a possible bias towards lighter handaxes in Group 1 (Emulation), and across groups in earlier attempts, masking any increased reduction by flaking as participants became more familiar with the process. Group 1 (Emulation) split and subsequently modified a broken piece of the porcelain blank in over half of all attempts, while all other groups only did so in less than a quarter of attempts (see table 5.1 for quantities). This could make the pieces, especially in Group 1 (Emulation), look artificially more reduced by flaking when they were actually smaller because of breakage and not because of more intense flaking.

5.3 Traditional Morphometric Framework

5.3.1 Knapping Intensity and Reduction

The first variables to be analysed will be those associated with knapping intensity and reduction, those being the weight lost from the handaxe during the knapping process, and the scar count and density, from flake negatives apparent on the handaxes.

5.3.1.1 Weight Lost

The summary statistics of weight lost during the knapping process of the handaxe blanks are outlined in Table 5.3. Several observations can be noted: firstly, and perhaps most notable is the coefficient of variation (CV) for the different groups. Measures of central tendency are key in analysing the distribution and spread of the range of morphological variation, and the CV in this case shows decreasing values (greater standardisation) correlating with simulated social learning methods which involve increased theory of mind capabilities. Similarly, the standard deviation (SD) of the groups decreases from Group 1 (Emulation) to Group 4 (Verbal), suggesting increased convergence of morphological range. Also of note is the increasing minimum amount of weight lost trending from Group 1 (Emulation) to the taught groups, showing increased reduction in the manufacturing of handaxes. Despite this, the maximum amount of weight lost in all groups is of similar values.

	Handaxe Total	Min (g)	Max (g)	Mean (g)	Median (g)	SD (g)	CV (%)
Weight lost	104	250.7	797.8	549.7	523.4	129.3	23.5
With BASM handaxes removed	73	250.7	797.8	496.2	484.2	106.5	21.5
Group 1 (Emulation)	25	252.9	772.6	602.4	656.9	146.7	24.4
Group 2 (Imitation)	26	250.7	769.3	543.1	552.8	129.7	23.9
Group 3 (Silent)	24	377.9	760.3	537.5	493.3	116.3	21.6
Group 4 (Verbal)	29	358.7	797.8	520.3	479.0	115.6	22.2

Table 5.3. Summary statistics for weight lost per group.

The weight lost from the handaxes in each group are shown in a box plot in Figure 5.1. The trend of decreasing variation from the emulation to the teaching groups is clear. The higher median weight lost in Group 1 (Emulation) and lower median weight lost in Group 4 (Verbal) was unexpected, as it was anticipated that flaking skills would be acquired faster in groups with more access to knowledge about the knapping process, and greater reduction would therefore be seen in Groups 2, 3 and 4, with the teaching groups (Groups 3 and 4) experiencing the highest reduction. Examination of the 'weight lost' variable through a Shapiro-Wilk normality test demonstrates that a normal distribution could not be assumed, and a Mann-Whitney pairwise test for significance was performed. The only groups which exhibited significant differences were between Groups 1 (Emulation) and 4 (Verbal), the most behaviourally contrasting methods in the experiment (*p*: 0.015).

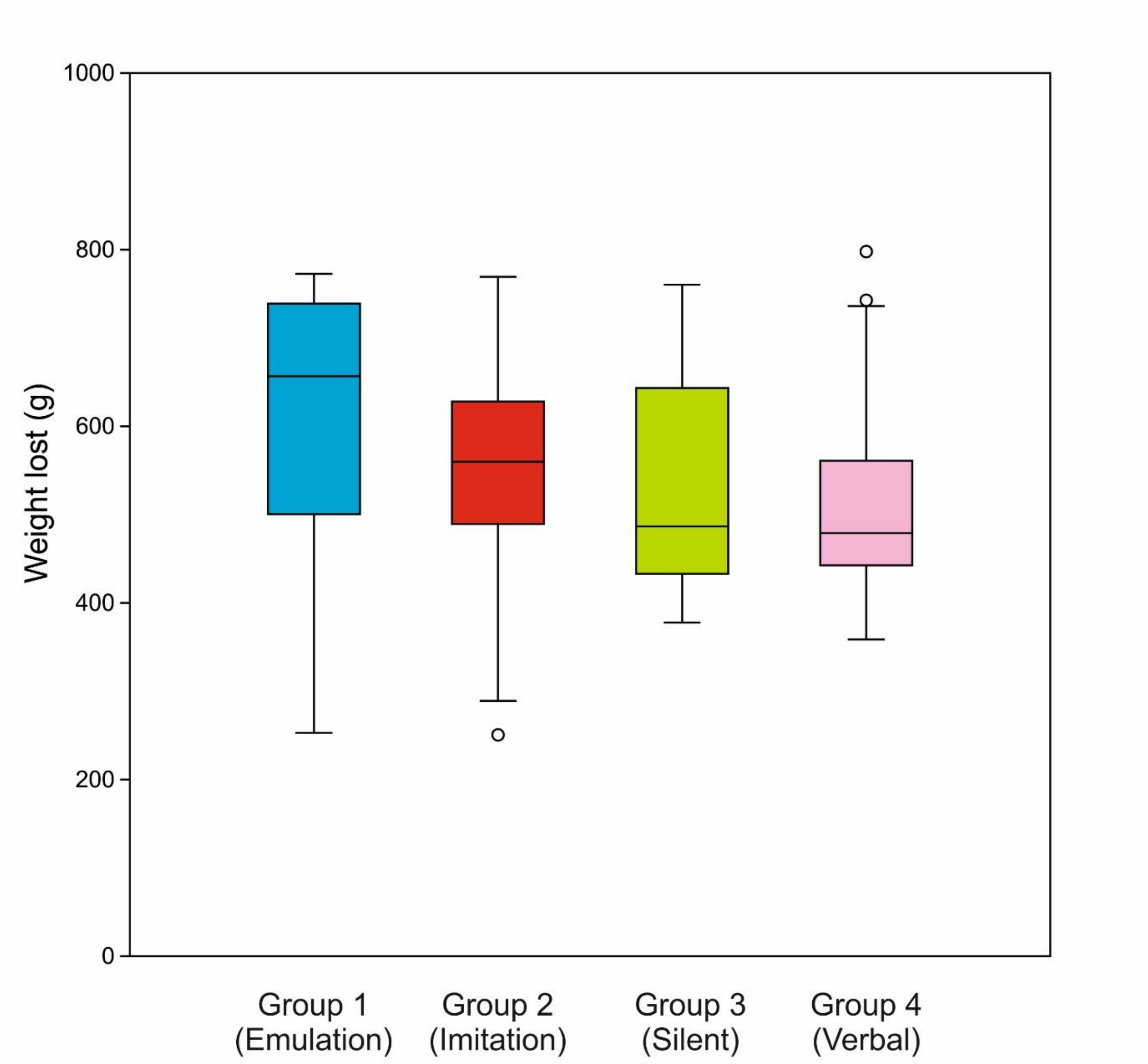


Figure 5.1. A box plot showing the weight lost per group

During the experiment, it was noted how often the participants, especially Group 1 (Emulation) accidentally broke their blank, and created a tool on the larger chunk that resulted from the breakage. This could influence why Group 1 (Emulation) shows much higher weight loss than the other groups. Therefore, the data was also analysed with the broken and subsequently modified (BASM) pieces removed, which accentuated the trend (Figure 5.2).

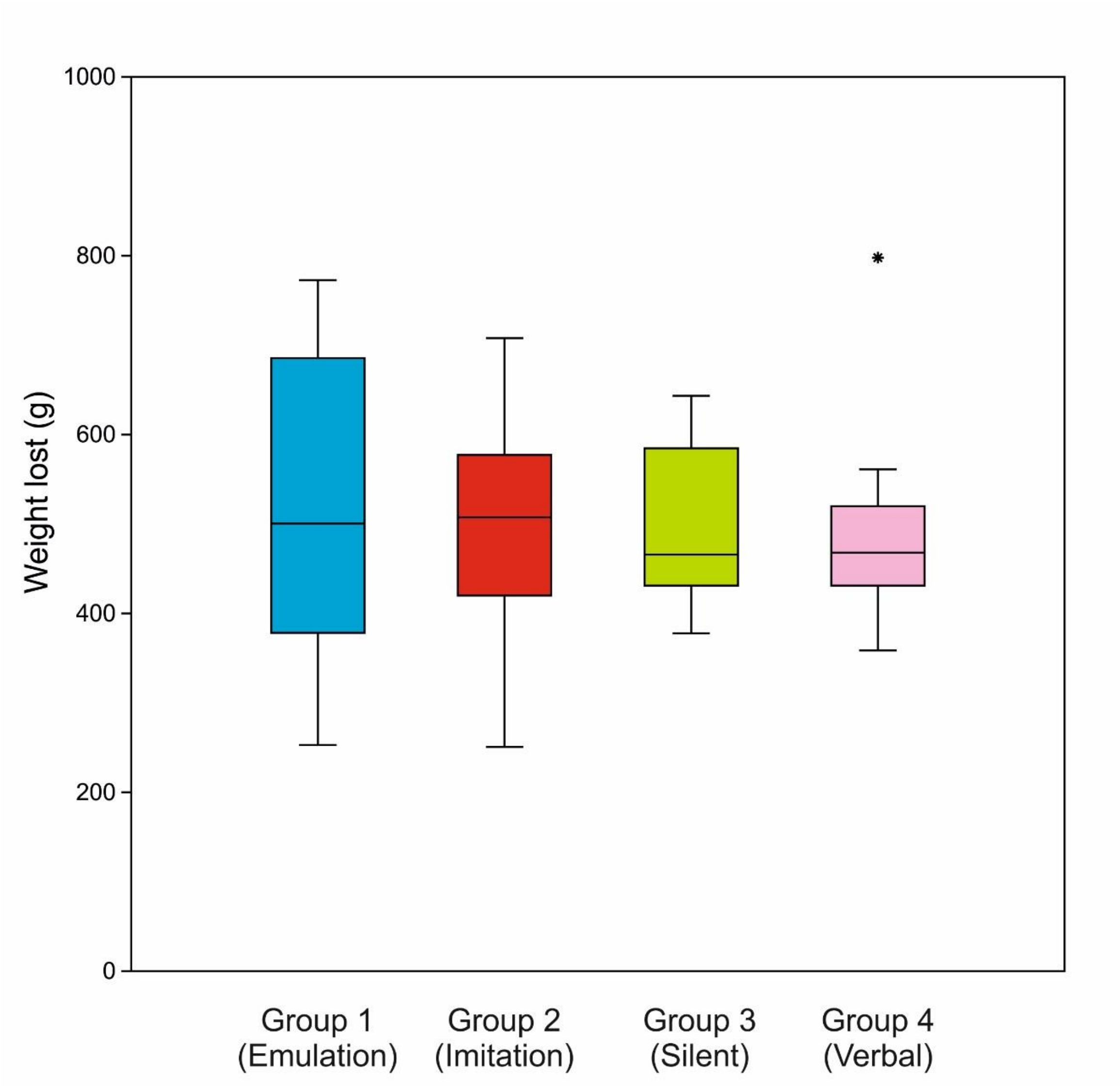


Figure 5.2. A box plot showing the weight lost per group, excluding BASM pieces.

When BASM handaxes are removed from the dataset, the boxplots for weight lost exhibit a different pattern. The median weight loss in all instances are of similar size (whereas in Figure 5.1, Group 1 (Emulation) had a much higher median weight loss value). While the range in weight lost did not reduce when BASM handaxes were removed in Group 1 (Emulation), the range did reduce among the other groups, decreasing in size throughout (most notably in the top groups). This suggests that the BASM handaxes, which were by necessity smaller than the handaxes produced on unbroken blanks, skewed the data towards higher weight lost values. This, however, was less pronounced in Group 1 (Emulation), which exhibited the same degree of variability even when this variable for increased variability was removed. Again, examination of the 'weight lost' variable through a Shapiro-Wilk normality test demonstrates that a normal distribution of examples excluding BASM handaxes could not be assumed, and a Mann-Whitney pairwise test for significance was performed. This test demonstrated that no groups were statistically different (see Chapter 6 for further discussion).

We can examine differences in weight lost as the experiment progressed when looking at the groups individually, separated by attempt number, (Figure 5.3). Generally, there is some reduction in the variation of weight lost with progressive attempts. This is most notable in Group 4 (Verbal), where the range of variability in weight lost becomes markedly low from attempt four onwards. The median throughout the attempts in Group 4 (Verbal) are remarkably consistent. Group 1 (Emulation) and 2 (Imitation) were the most erratic in their variability, both featuring varying medians of weight lost and ranges of variability. Group 3 (Silent) falls somewhere in between these two distinctions, with some degree of consistency in the median weight lost, as well as a converging range of variability through the attempts.

This analysis was not repeated for BASMs due to issues of sample size; whereas the smallest number featured in groups' attempt with BASMs included is three (but most often 4 or 5), with BAMS discluded group size could be as low as 1.

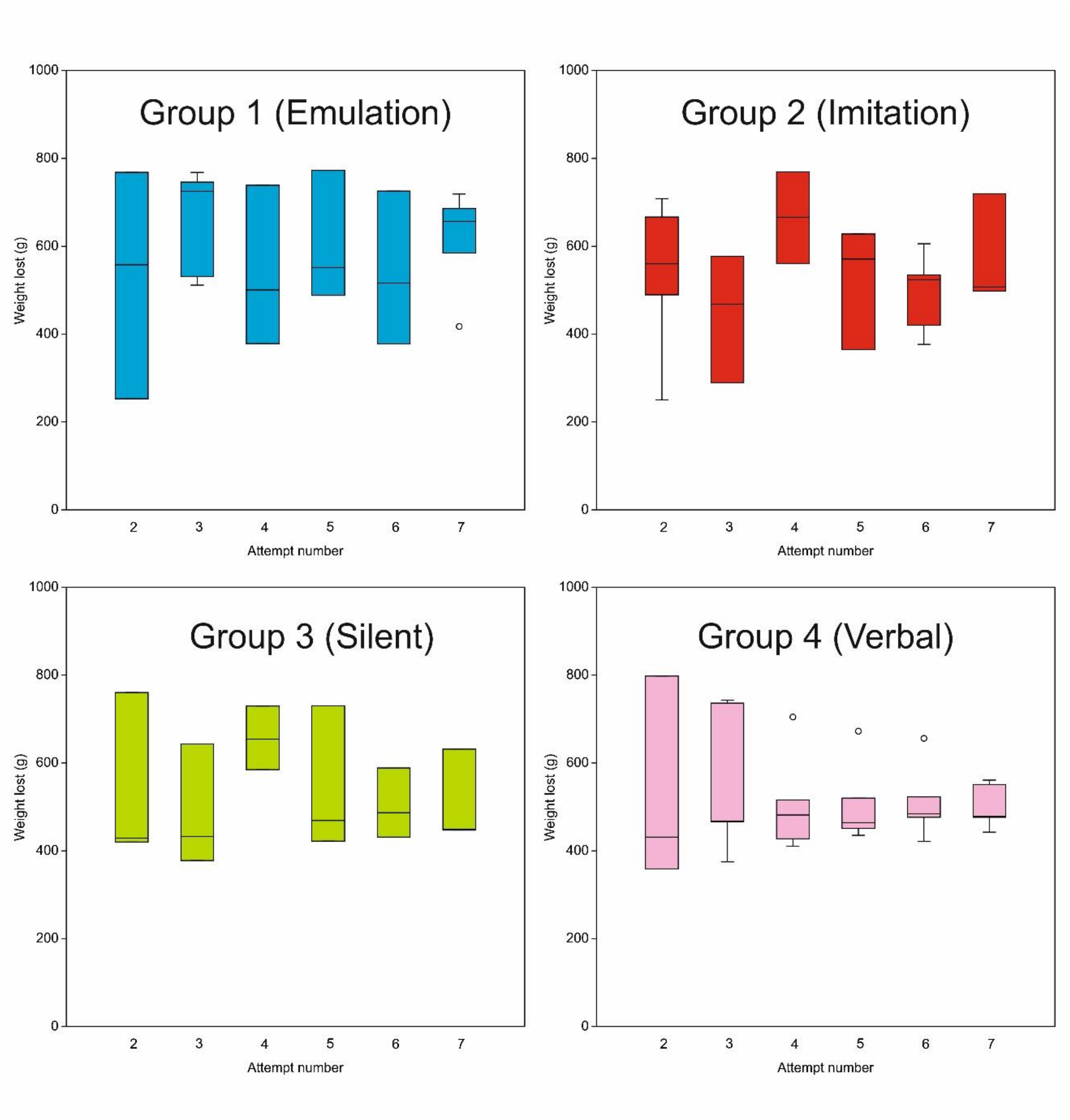


Figure 5.3. Box plots of the weight lost for each attempt's handaxe in the different groups.

5.3.1.2 Scar Count and Scar Density

The summary statistics for the number of scars present on each handaxe (over 2cm in axial length), overall and for each group can be viewed in Table 5.4. Similar to the trend in the coefficient of variation seen in weight lost, there is a decrease in values from Group 1 (Emulation) to Group 4 (Verbal), suggesting increasing standardisation correlating with increasingly complex (simulated) social learning methods. The standard deviation, median and mean for Groups 2-4 are all relatively similar but stand in stark contrast to Group 1 (Emulation). Reasons for this will be discussed below in regards to the removal of BASM handaxes.

Figure 5.4 shows a box plot for the scar counts present in each group's handaxes. Group 1 (Emulation) shows the lowest interquartile range, in addition to the lowest median count. In this measurement, there is not a reduction in the range of variation of scar count in the higher fidelity social learning groups, and in fact Group 1 has the smallest range for these values. With no explicit coaching, it is not surprising that Group 1 (Emulation) created handaxes with less flake scars left on the surface than the other groups. It was noticed during the data collection that many of the scars on the handaxes in Group 1 (Emulation) were small and under the 2cm threshold to be counted (and battering and repeated step fractures were common). Anecdotally, the other groups seemed to detach larger flakes, and that will be presented in the section to follow (debitage analysis), when flakes between 2-3cm and greater than 3cm were catalogued as a measurement of participants gaining skill in detaching larger flakes.

	Handaxe Total	Min (n =)	Max (n =)	Mean (n =)	Median (n =)	SD (n =)	CV (%)
Overall	104	4.0	24.0	13.5	13.0	4.4	32.2
Group 1 (Emulation)	25	5.0	18.0	10.6	10.0	3.5	32.7
Group 2 (Imitation)	26	4.0	21.0	13.5	13.0	4.3	32.1
Group 3 (Silent)	24	7.0	24.0	15.7	15.5	4.1	26.1
Group 4 (Verbal)	29	8.0	23.0	14.3	13.0	4.1	28.9

Table 5.4. Summary statistics for scar count present on each handaxe.

A Shapiro-Wilk test showed that all four groups feature a normal distribution, and a pairwise test for equal means (two-sample t test) was performed (Table 5.5) to determine statistical significance. Group 1 (Emulation) was statistically different from all other simulated social learning groups, but the other three groups could not be statistically distinguished.

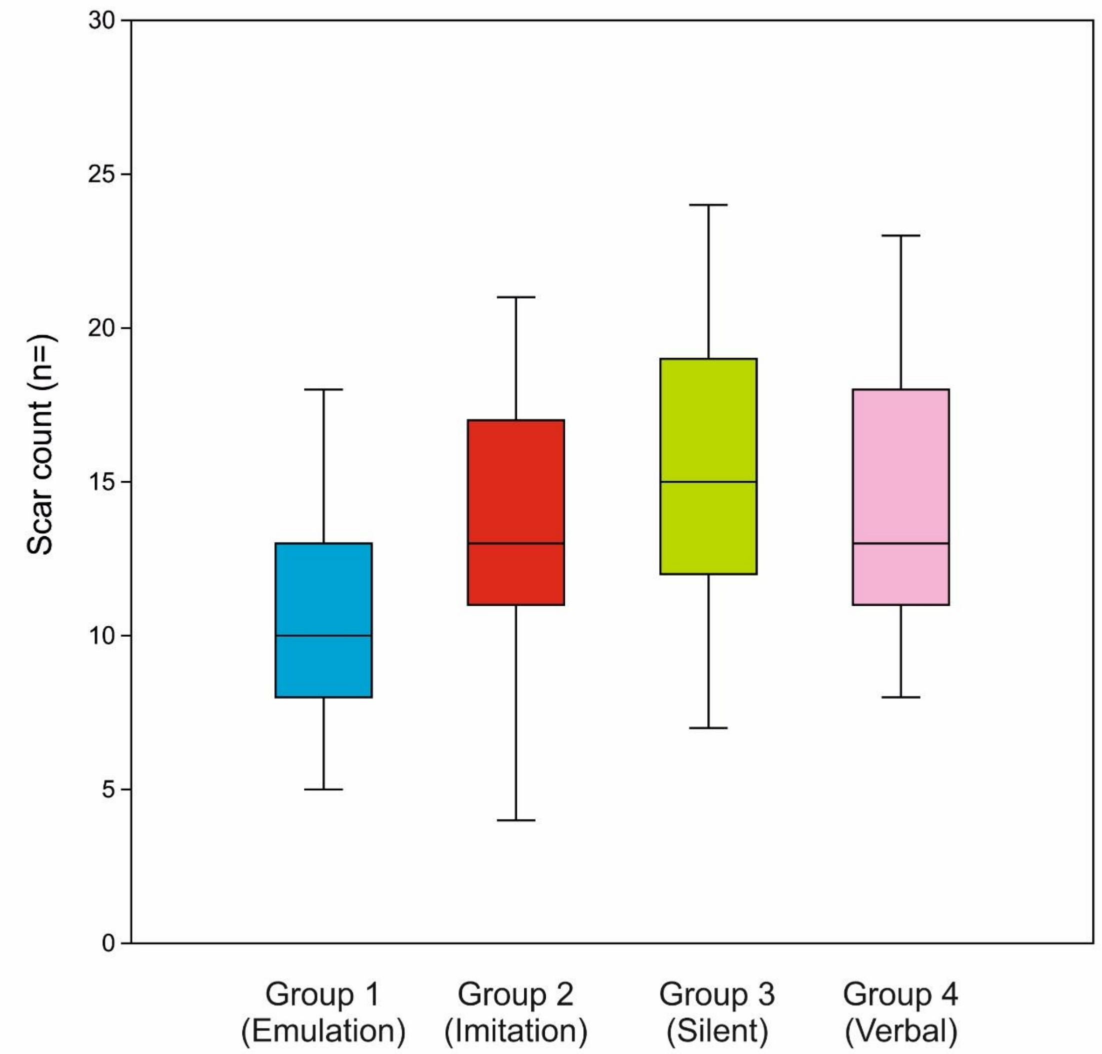


Figure 5.4. A box plot of scar count (larger than 2cm in axial length) for handaxes in each group.

	Group 1 (Emulation)	Group 2 (Imitation)	Group 3 (Silent)	Group 4 (Verbal)
Group 1 (Emulation)		***p* : 0.013**	***p* : < 0.001**	***p* : < 0.001**
Group 2 (Imitation)	***p* : 0.013**		*p* : 0.076	*p* : 0.487
Group 3 (Silent)	***p* : < 0.001**	*p* : 0.076		*p* : 0.248
Group 4 (Verbal)	***p* : < 0.001**	*p* : 0.487	*p* : 0.248	

Table 5.5 Pairwise tests for equal means of scar count (Monte Carlo simulations, N = 9999). Significance to 95% confidence (α: 0.05).

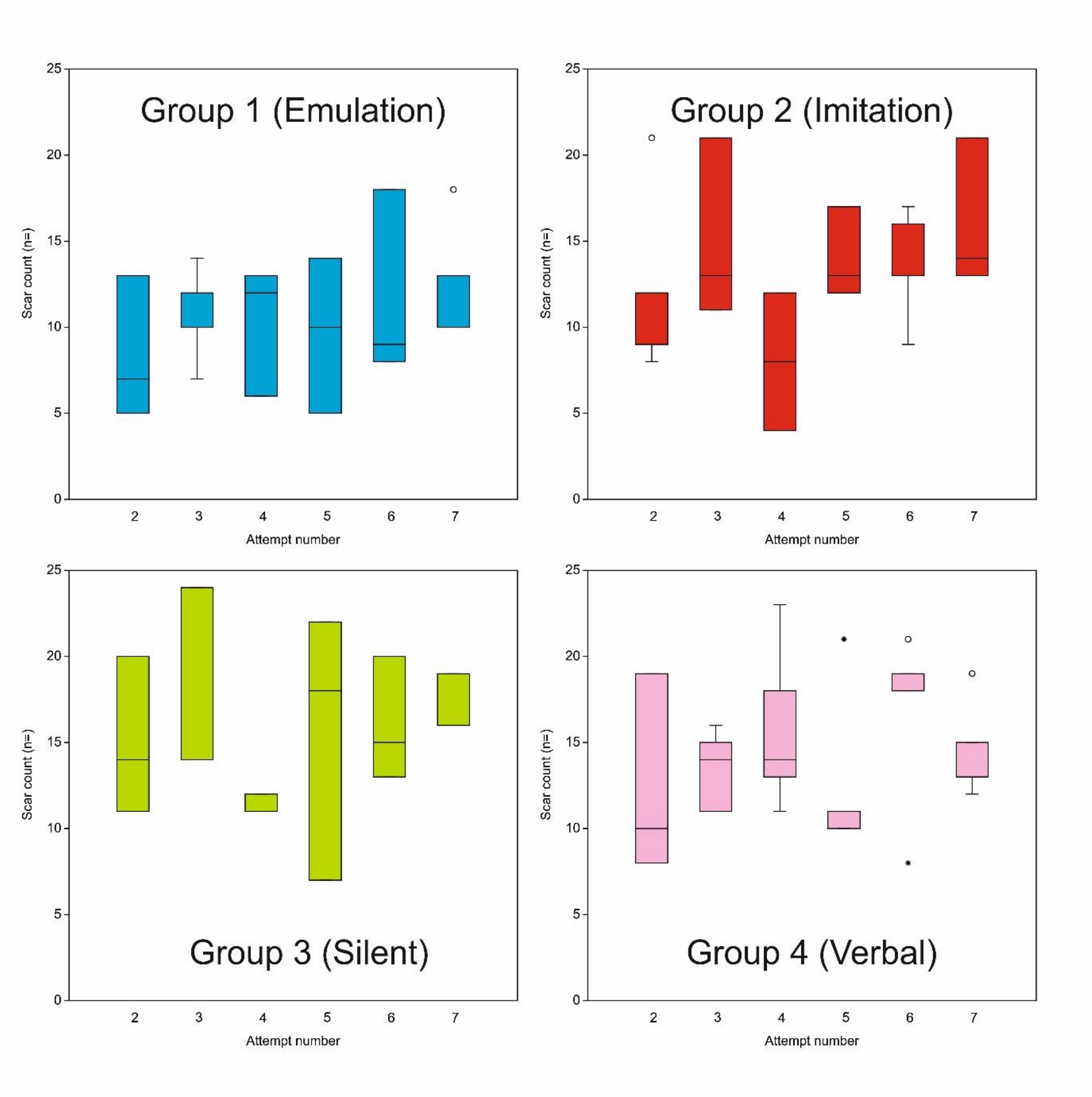


Figure 5.5. Box plots of scar count for each attempt's handaxe in the different groups.

By looking at the data for scar count separated by attempt number, we can observe general trends and patterns across groups as participants in each group progressed through the experiment (Figure 5.5).

The graphs do not appear to show the decreasing range of variability as seen in the weight lost measurement, for example, but there does appear to still be somewhat of a trend. Group 4 (Verbal) does demonstrate this marked reduction in the range of variability of scar count. One reason that the variability does not appear to reduce to the extent seen in the previous variable could be the issue of broken and subsequently modified (BASM) handaxes, where some group's pieces broke in half, making much smaller tools. This affects the amount of surface that can retain flake scars. It biases the groups with smaller tools that feature fewer scars greater than 2cm, where the actual flake scar density may be the same or similar to handaxes produced on unbroken blanks.

In order to investigate the number of scars in relation to the size of the piece, instead of removing the BASM handaxes, a measurement of scar density was created, by taking the measure of scar count and dividing it by the weight of the handaxe. This gives a measurement of the density of scars over 2cm in size on a handaxe. A higher number reflects a higher average density of scars on the tool surface. Summary statistics for this measure are presented in Table 5.6. The coefficient of variation for this measure does not decrease consistently from Group 1 (Emulation) to Group 4 (Verbal) as it does with weight lost and scar count. As can be seen in Figure 5.6, a number of outliers particularly in Group 4 (Verbal) can be observed. Because of the small dataset, outliers such as these have a large effect.

	Handaxe Total	Min (n =)	Max (n =)	Mean (n =)	Median (n =)	SD (n =)	CV (%)
Overall	104	0.011	0.121	0.046	0.044	0.019	41.377
Group 1 (Emulation)	25	0.011	0.089	0.046	0.043	0.020	42.940
Group 2 (Imitation)	26	0.013	0.108	0.045	0.046	0.020	46.165
Group 3 (Silent)	24	0.029	0.092	0.049	0.048	0.015	29.859
Group 4 (Verbal)	29	0.018	0.122	0.045	0.040	0.020	46.468

Table 5.6. Summary statistics for the scar density of each handaxe.

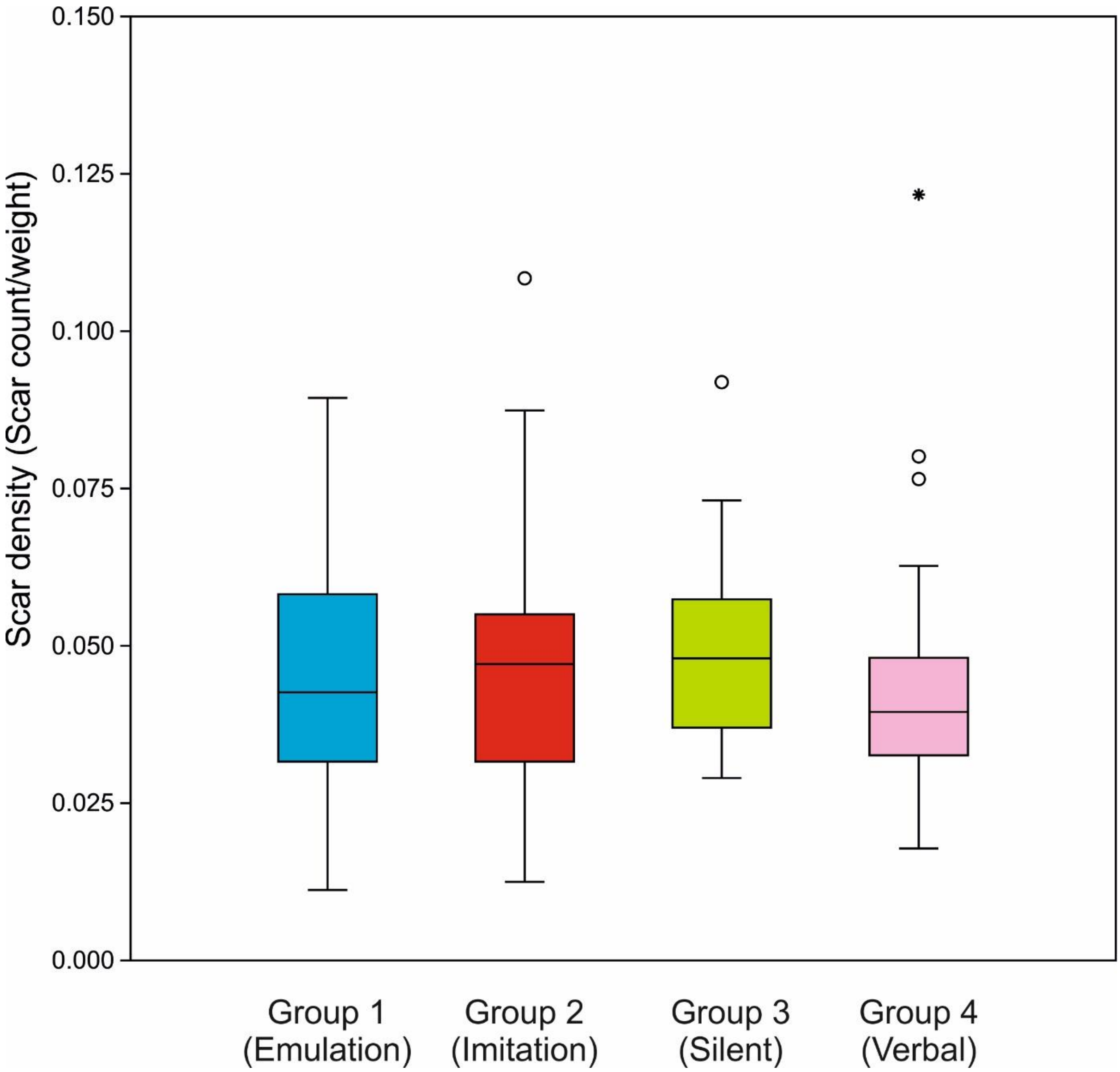


Figure 5.6. A box plot of scar density for handaxes in each group.

In Figure 5.6, the scar density measure shows a more expected range of variability, excluding outliers. A non-normal distribution was concluded through a Shapiro-Wilk test for normality; therefore, a Mann-Whitney pairwise test for equal medians was performed. No statistical significance was found between any of the four simulated social learning groups. What this suggests is that in the measure of scar count, where Group 1 (Emulation) was statistically different from all groups (but not by the measure of scar density), scar density removed the issue introduced by the BASM handaxes.

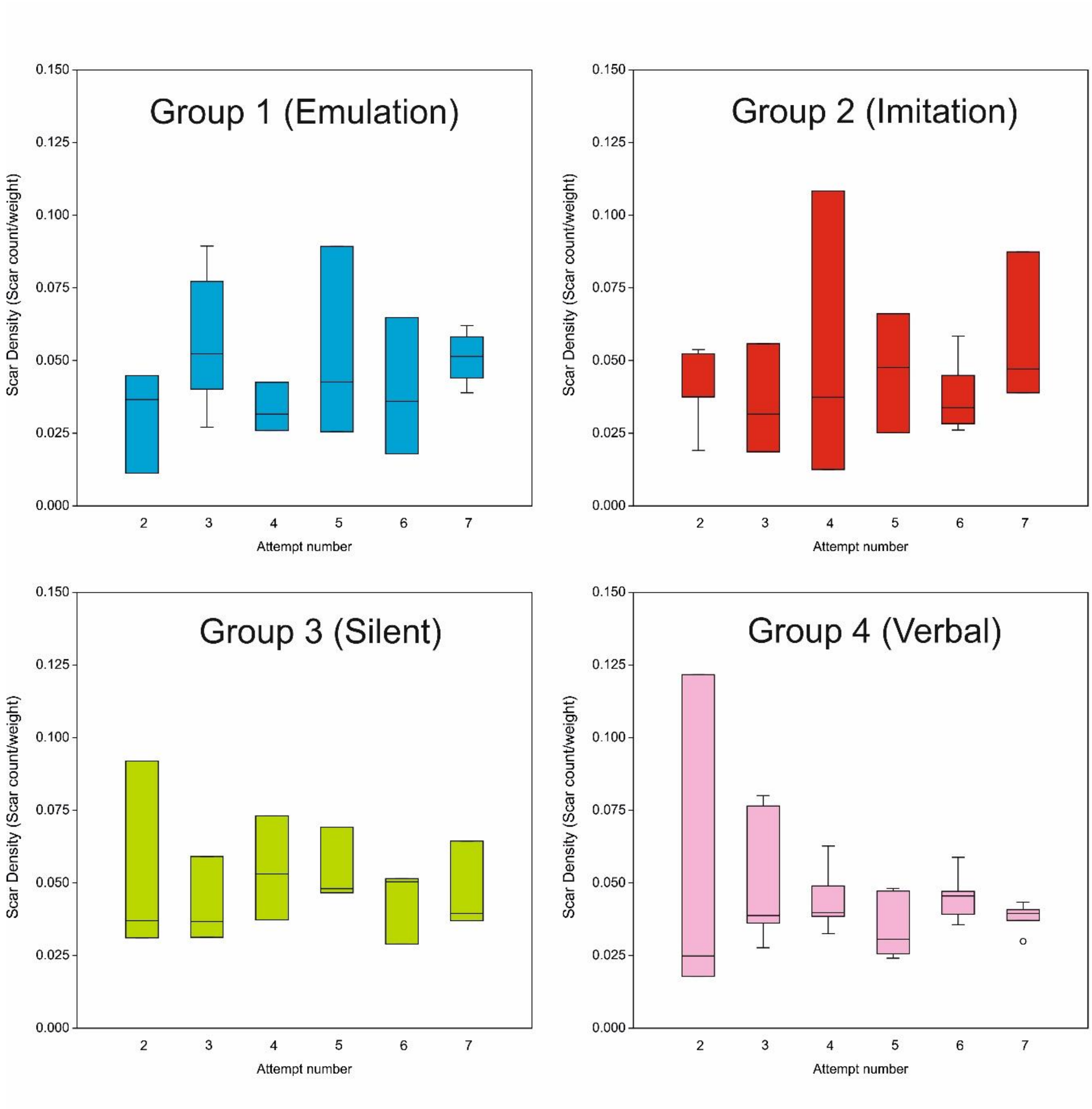


Figure 5.7. Box plots of scar density for each attempt's handaxe in the different groups.

It is interesting to note that when broken down by group to analyse scar density of the handaxes, that the taught groups begin with the highest levels of variability and quickly descend to lower ranges of variability (Figure 5.7). One possibility for the large initial ranges could be that flaking skill is acquired earlier by some members of the taught groups, which increases the potential range of variability within these groups. Group 1 (Emulation) and Group 2 (Imitation) feature less or no coaching on successful flaking, and so all participants in these groups have very low scar count evident on the pieces, which results in a low range of variability. In Group 2 (Imitation), an increase in variability becomes apparent by attempt 4, which could demonstrate this slower knowledge acquisition by some participants of this group, producing more potential for flaking success after a few initial attempts.

5.3.2 Flake Attributes

The second set of variables to be analysed using traditional measurements will be aspects in the range of flake variation between groups, that of the flake count produced per handaxe and of the weight of the flakes produced in the creation of each handaxe.

5.3.2.1 Flake count

As noted in Chapter Four, flakes larger than 2cm in axial length were included in the flake count analyses. The summary statistics for flake count are provided in Table 5.7. This measure presents an expected trend of increasing standardisation correlating with simulated social learning methods which require increasing levels of theory of mind. This is expressed through both the coefficient of variation and the standard deviation, which decrease in values from Group 1 (Emulation) through to the taught groups. Notable is the low minimum flake count in Group 1 (Emulation) and the high minimum flake count in the taught groups, possibly a direct result of Group 1 having many more BASM handaxes, which effects the number of successful detached flakes from a handaxe.

	Handaxe Total	Min (n =)	Max (n =)	Mean (n =)	Median (n =)	SD (n =)	CV (%)
Overall	104	6.0	73.0	40.2	40.0	14.4	35.9
With BASM handaxes removed	73	9.0	73.0	43.5	43.0	13.3	30.5
Group 1 (Emulation)	25	6.0	51.0	28.9	28.0	13.2	45.6
Group 2 (Imitation)	26	12.0	63.0	39.5	39.5	13.4	33.8
Group 3 (Silent)	24	23.0	73.0	49.3	49.5	15.1	30.7
Group 4 (Verbal)	29	26.0	63.0	43.2	43.0	8.7	20.1

Table 5.7. Summary statistics for total flake count for each handaxe.

When analysing the total flake count per handaxe for each group (Figure 5.8), a repeated trend emerges. Group 4 (Verbal) exhibits the most convergent and smallest range of variability. We actually see in this measure that Group 1 (Emulation) and Group 3 (Silent) have similar ranges of variability, albeit with differing medians. There is a general trend of higher flake count per handaxe correlated with increased knowledge imparted in the simulated social learning method. Group 4 (Verbal) has a slightly lower median flake count value than Group 3 (Silent) however due to their similar access to knowledge about the flaking task (in comparisons to Group 1 and 2) it is not expected that Group 4 (Verbal) would significantly outperform Group 3, with greater flaking. A Shapiro-Wilk normality test confirms that both groups demonstrate a normal distribution, and a two-sample t test was performed to test for significance between these two groups. Trending significance was documented (t: 1.810, p: 0.077), however no statistical significance (to 95% confidence) was identified.

Further testing through a Shapiro-Wilk test for normality showed a normal distribution for all groups, and a pairwise two-sample t test for equal means was conducted (Table 5.8).

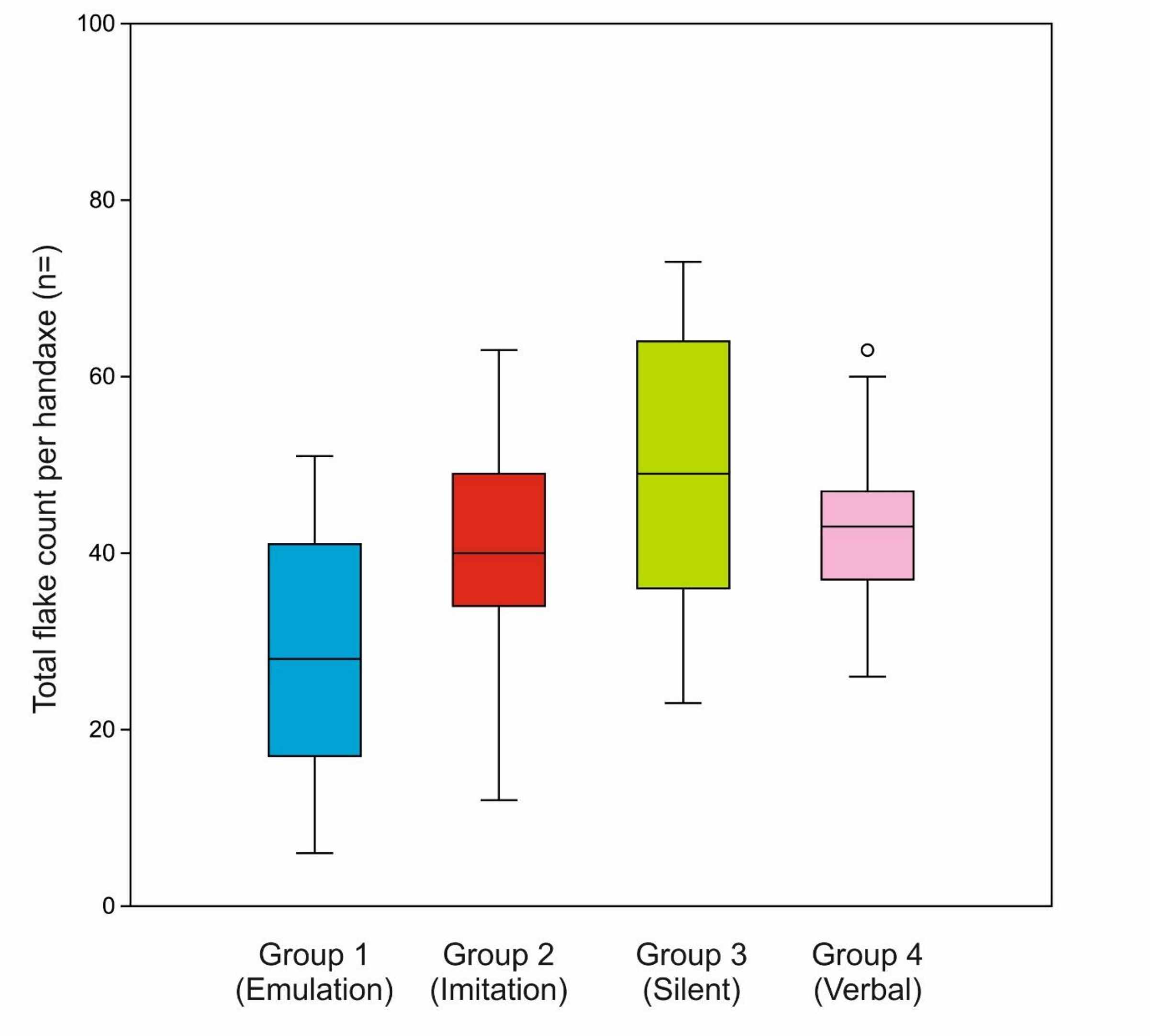


Figure 5.8. A box plot of total flake count for handaxes in each group.

	Group 1 (Emulation)	Group 2 (Imitation)	Group 3 (Silent)	Group 4 (Verbal)
Group 1 (Emulation)		***p* : 0.006**	***p* : < 0.001**	***p* : < 0.001**
Group 2 (Imitation)	***p* : 0.006**		***p* : 0.018**	*p* : 0.221
Group 3 (Silent)	***p* : < 0.001**	***p* : 0.018**		*p* : 0.078
Group 4 (Verbal)	***p* : < 0.001**	*p* : 0.221	*p* : 0.078	

Table 5.8. Pairwise tests for equal means of flake count (Monte Carlo simulations, N = 9999). Significance to 95% confidence (α: 0.05).

Similarly to the pairwise two-sample *t* test conducted in the previous section, Group 1 (Emulation) was shown to be statistically different to all other groups. The significance between other groups varied.

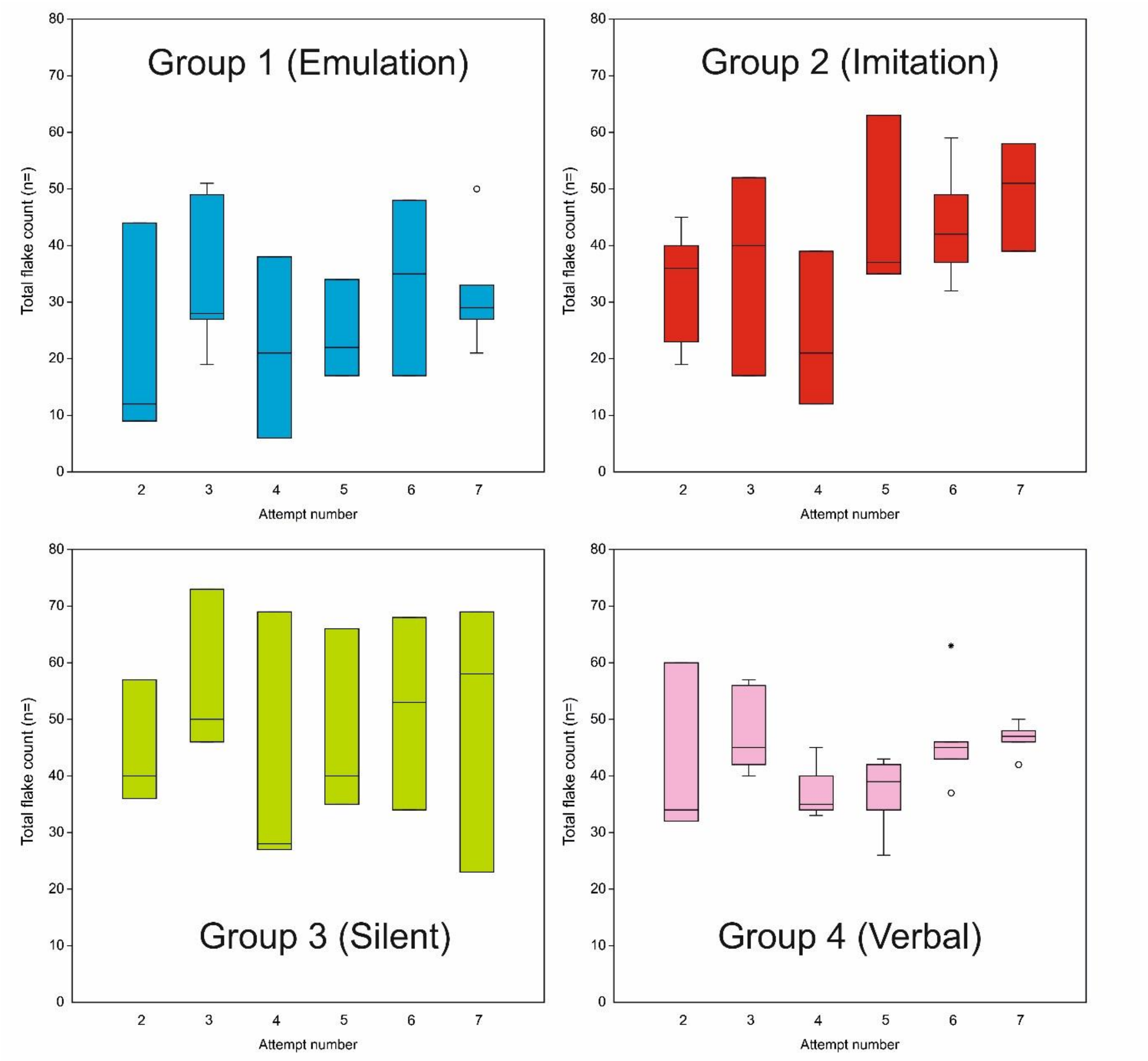


Figure 5.9. Box plots of total flake count for each attempt's handaxe in the different groups.

In Figure 5.9, it can be observed that the median total flake count increases as participants progressed through the experiment, as their flaking skill increased and they made progressively more flakes when replicating the handaxes. In Group 4 (Verbal) it can be seen again see a very marked convergence in the range of total number of flakes produced per handaxe, and to some extent in Group 1 (Emulation) and Group 2 (Imitation). The most contrasting difference between Group 1 and 2 is the much lower number of flakes produced in Group 1 (Emulation)'s handaxes. Interestingly, the converging range of variability is not observed in Group 3 (Silent).

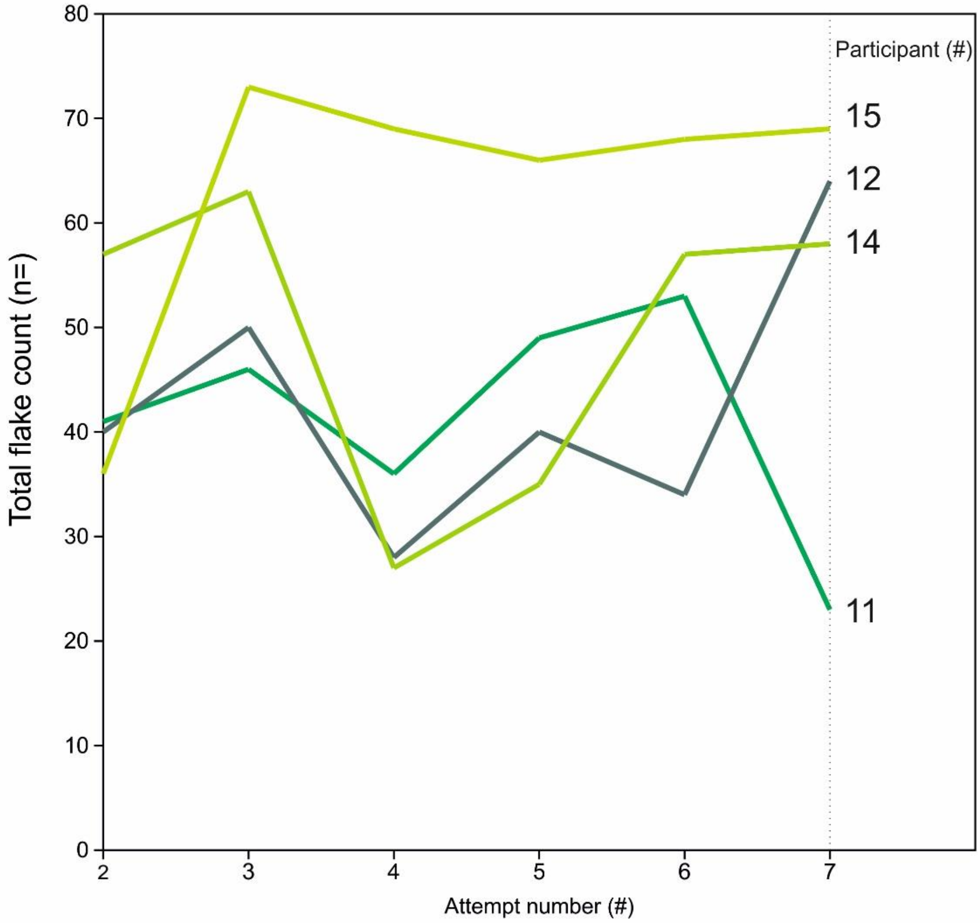


Figure 5.10. A line graph showing total flake count for each participant throughout the experiment in Group 3 (Silent).

In Figure 5.10, it is observed that the total flake count for each participant in Group 3 (note: participant 13 was excluded from the data analysis, see Section 5.1). A number of reasons for the high range of variability exhibited can be identified.

Participant 11's final handaxe exhibits a very low flake count, most likely due to the fact that the handaxe split at the end of their attempt, resulting in preventing further refinement and further flakes being removed. Participant 15 had a consistently high flake count throughout the experiment, from attempt 3 onwards, resulting in a widened range of variability for the box plots in Figure 5.9. Finally, three of the four handaxes in Attempt 4 were either BASM handaxes or split on the final blow, preventing further refinement and flaking, resulting in this attempt's low flake count. These variables, coupled with a small dataset, explain the higher range of variability exhibited in Group 3 (Silent)'s total flake count.

In Figure 5.11, where BASM handaxes have been removed from the sample, there is a slight increase in the range of variability exhibited in Group 4, but with the general pattern of increasing flake count between Groups 1-3 remaining the same. The analysis was not further analysed through attempt number due to issues of sample size.

A Shapiro-Wilk test for normality concluded that all four groups exhibited a normal distribution, and a pairwise two sample *t* test for equal means was performed (Table 5.9). Analyses were similar to the previous analyses incorporating the BASM handaxes, in that Group 1 (Emulation) could be discriminated from all other groups. Analyses however differed in that Group 2 (Imitation) and Group 3 (Silent) were statistically significant to a 95% confidence level, but the general pattern observed through previous analyses was maintained.

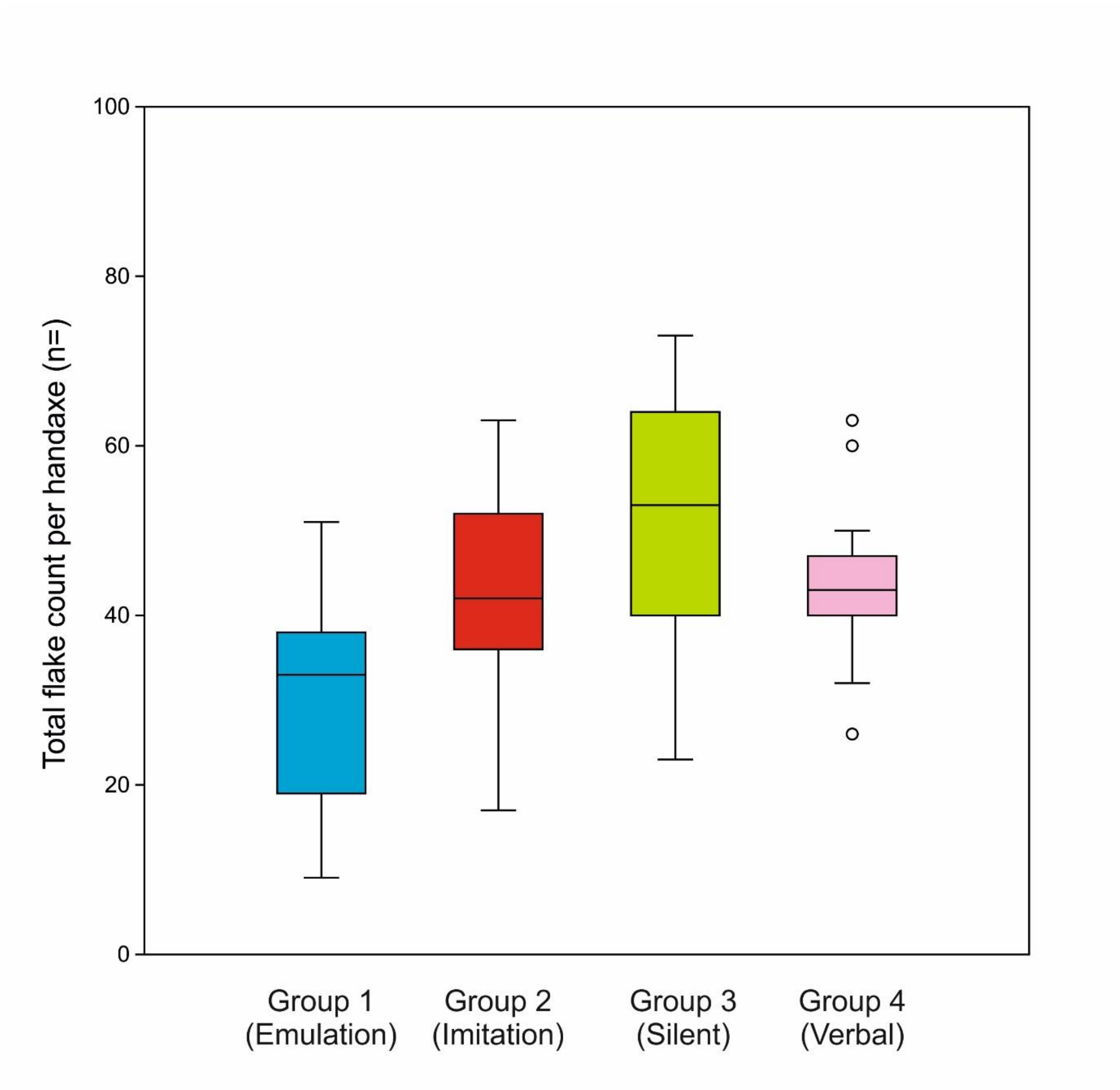


Figure 5.11. A box plot showing the total flake count per handaxe excluding BASM pieces.

	Group 1 (Emulation)	Group 2 (Imitation)	Group 3 (Silent)	Group 4 (Verbal)
Group 1 (Emulation)		***p* : 0.010**	***p* : < 0.001**	***p* : 0.002**
Group 2 (Imitation)	***p* : 0.010**		*p* : 0.057	*p* : 0.912
Group 3 (Silent)	***p* : < 0.001**	*p* : 0.057		***p* : 0.032**
Group 4 (Verbal)	***p* : 0.002**	*p* : 0.912	***p* : 0.032**	

Table 5.9. Pairwise tests for equal means of flake count excluding BASM handaxes (Monte Carlo simulations, N = 9999). Significance to 95% confidence (α: 0.05).

5.3.2.2 Flake Weight

As noted in Chapter Four, flakes larger than 2cm in axial length were included in the flake weight analyses. Summary statistics for flake weight are shown in Table 5.10. The coefficient of variation values further demonstrated a trend from the largest range of variation in Group 1 (Emulation), to the smallest range of variation in Group 4 (Verbal), suggesting increased morphological standardisation correlating to groups with simulated social learning methods that incorporated theory of mind capability. The mean and median values show a trend where Group 1 (Emulation) exhibited the lowest average flake weights for each handaxe, and the taught groups exhibited the highest, suggesting in conjunction with the minimum values that the taught groups' handaxes not only converged more, but that the participants acquired flaking knowledge more quickly.

	Handaxe Total	Min (g)	Max (g)	Mean (g)	Median (g)	SD (g)	CV (%)
Overall	104	36.7	442.3	214.1	207.2	86.8	40.5
With BASM handaxes removed	73	36.7	442.3	226.6	222.1	85.0	37.5
Group 1 (Emulation)	25	45.2	340.2	160.3	143.3	77.9	48.6
Group 2 (Imitation)	26	36.7	320.0	201.8	197.9	79.9	39.6
Group 3 (Silent)	24	127.9	442.3	260.3	234.8	89.9	34.5
Group 4 (Verbal)	29	124.3	391.3	233.2	219.4	72.4	31.1

Table 5.10. Summary statistics for total flake weight for each handaxe.

The box plot in Figure 5.12 shows flake weight in a visual representation, demonstrating many of the relationships observed in the above summary statistics and in Figure 5.11. The relationship between flake weight and flake count, for pieces bigger than 2cm, will be further discussed in the following section.

A Shapiro-Wilk test for normal distribution was conducted, which showed that a normal distribution could not be assumed for Group 3 (Shapiro-Wilk: 0.913, *p*: 0.04). Therefore, a pairwise Mann-Whitney test for equal medians was performed (Table 5.11).

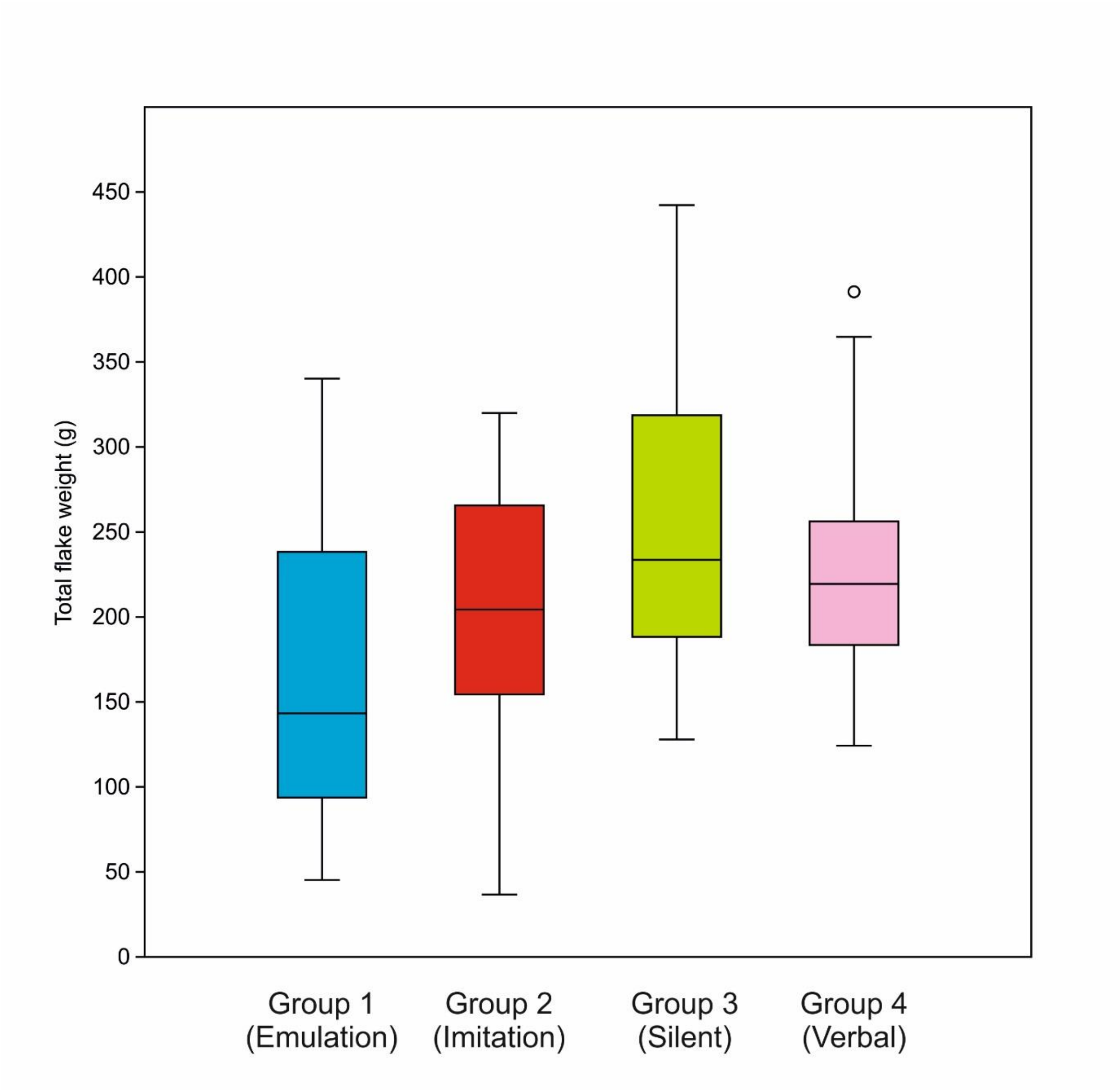


Figure 5.12. A box plot of total flake weights for handaxes in each group.

	Group 1 (Emulation)	Group 2 (Imitation)	Group 3 (Silent)	Group 4 (Verbal)
Group 1 (Emulation)		*p* : 0.053	***p* : < 0.001**	***p* : 0.001**
Group 2 (Imitation)	*p* : 0.053		*p* : 0.067	*p* : 0.266
Group 3 (Silent)	***p* : < 0.001**	*p* : 0.067		*p* : 0.257
Group 4 (Verbal)	***p* : 0.001**	*p* : 0.266	*p* : 0.257	

Table 5.11. Pairwise tests for equal means of flake weight for the different groups (Monte Carlo simulations, N = 9999). Significance to 95% confidence (α: 0.05).

The table shows that statistically significant differences to 95% confidence were shown in relationship to Group 1 (Emulation) with the taught groups, and trending significance noted between Group 1 (Emulation) and Group 2 (Imitation).

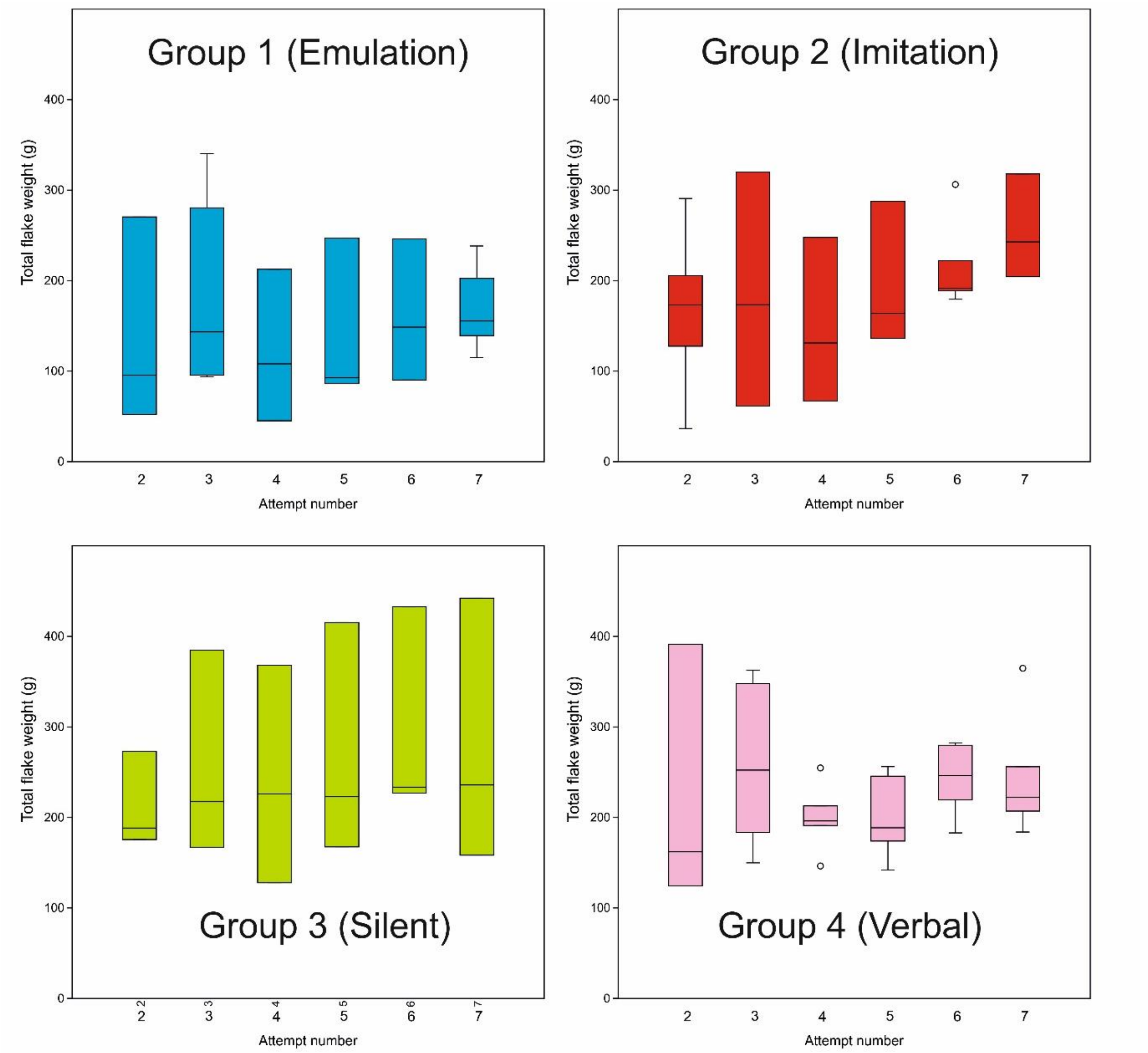


Figure 5.13 Box plots of total flake weight for each attempt's handaxe in the different groups.

Figure 5.13 shows a very convergent pattern for Group 4 (Verbal), echoing many of the figures seen previously, however in this case as in the measure of flake count, Group 3 (Silent) actually increases in variability. Reasons for this may be explained through BASM handaxes (see below), and will be elaborated on in Chapter Six. The trends and relationship seen in Figure 5.14 are broadly similar to those seen in flake count (Figure 5.9), suggesting a strong relationship between flake weight and flake count. This will be further discussed in Section 5.3.2.3.

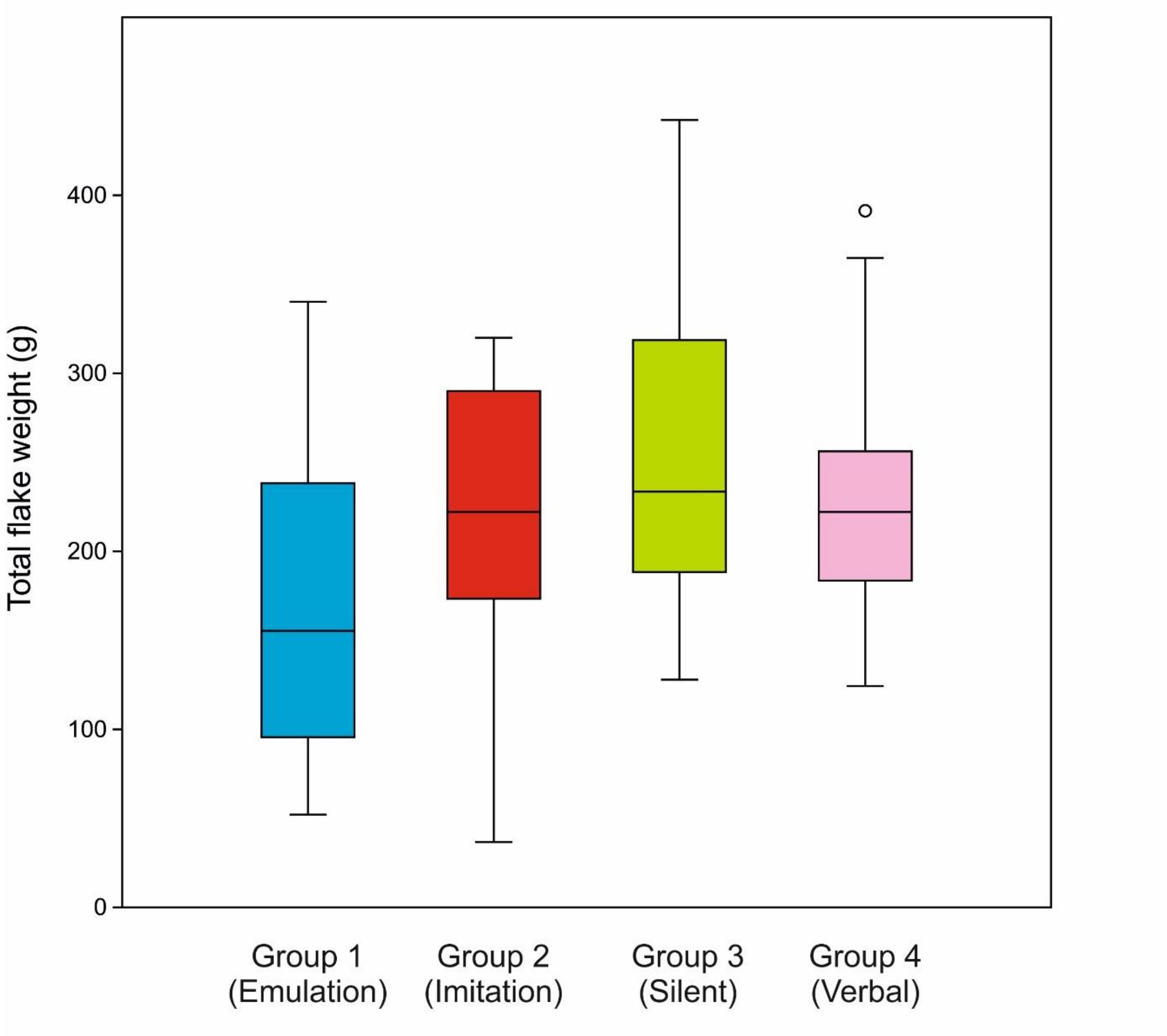


Figure 5.14. A box plot showing the total flake weight per handaxe excluding BASM pieces.

When BASM handaxes are removed from the analysis, it appears that no difference is observed, with respect to the range of variation, median total flake weights, and their interquartile ranges. A Shapiro-Wilk test for normality again highlighted that Group 3 (Silent) could not be assumed to have a normal distribution (Shapiro Wilk:

0.902, *p*: 0.047). Therefore, a pairwise Mann-Whitney test for equal medians was performed and it was confirmed that the exclusion of BASM handaxes had little to no effect on the statistical significance scores produced (Table 5.12).

	Group 1 (Emulation)	Group 2 (Imitation)	Group 3 (Silent)	Group 4 (Verbal)
Group 1 (Emulation)		*p* : 0.073	***p* : 0.016**	***p* : 0.027**
Group 2 (Imitation)	*p* : 0.073		*p* : 0.491	*p* : 0.961
Group 3 (Silent)	***p* : 0.016**	*p* : 0.491		*p* : 0.363
Group 4 (Verbal)	***p* : 0.027**	*p* : 0.266	*p* : 0.363	

Table 5.12. Pairwise tests for equal means of flake weight (excluding BASM handaxes) for the different groups (Monte Carlo simulations, N = 9999). Significance to 95% confidence (α: 0.05).

5.3.3 Summary of Results

Likely the most useful measure for quantifying the range of morphological variability feature in each of the categories above is the coefficient of variation value for each group (Figure 5.15). When synthesised, it provides a broad picture where a general trend of larger variation in Group 1 (Emulation) trends to more standardised CV values in the taught simulated social learning groups. The main exception to this is the scar density measure, where the hypothesised trend was not exhibited. As the measure for incorporating flake scars was above 2cm, the increased scar density may not be visible given the decreasing size of handaxes produced. Therefore, it is unknown would have shown a result consistent with other measures if flake scars above 10mm were also considered. In all other measures featured the expected trend from low fidelity social learning and high morphological variability, to high fidelity social learning and low morphological variability can be observed in a number of independent variables.

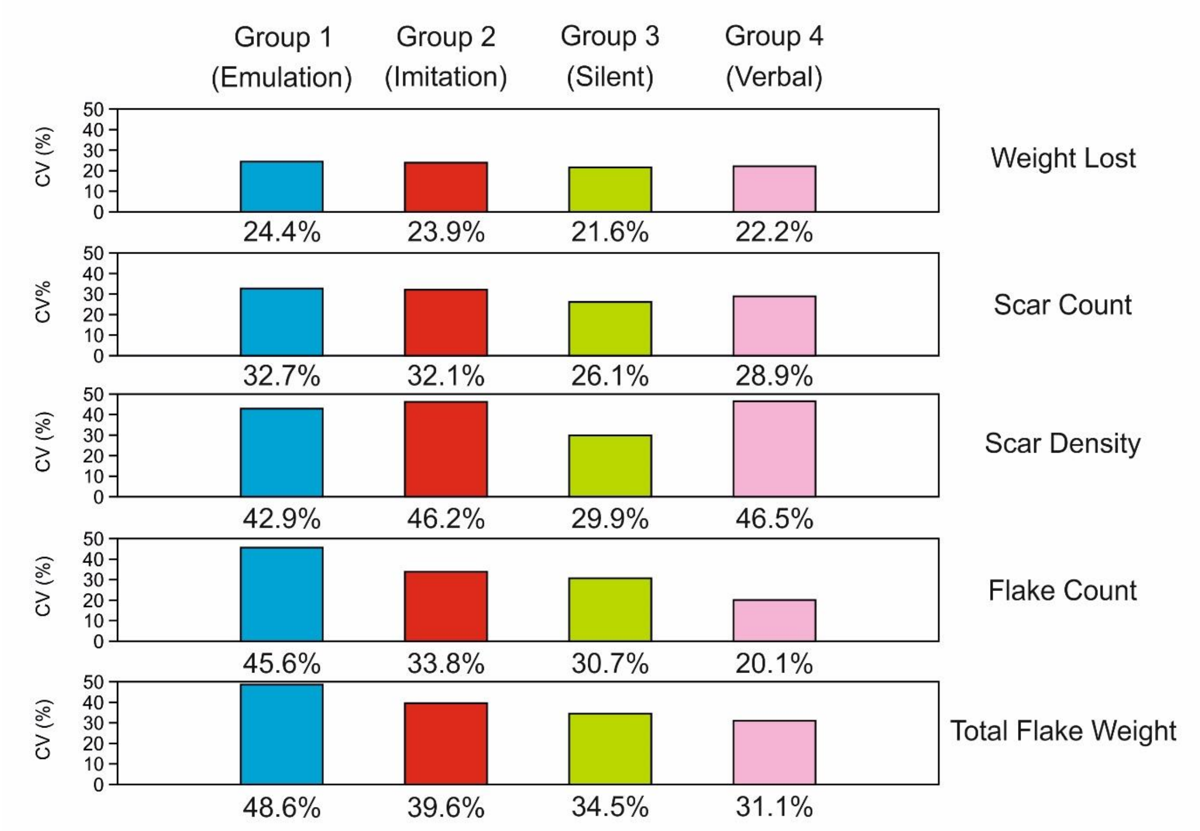


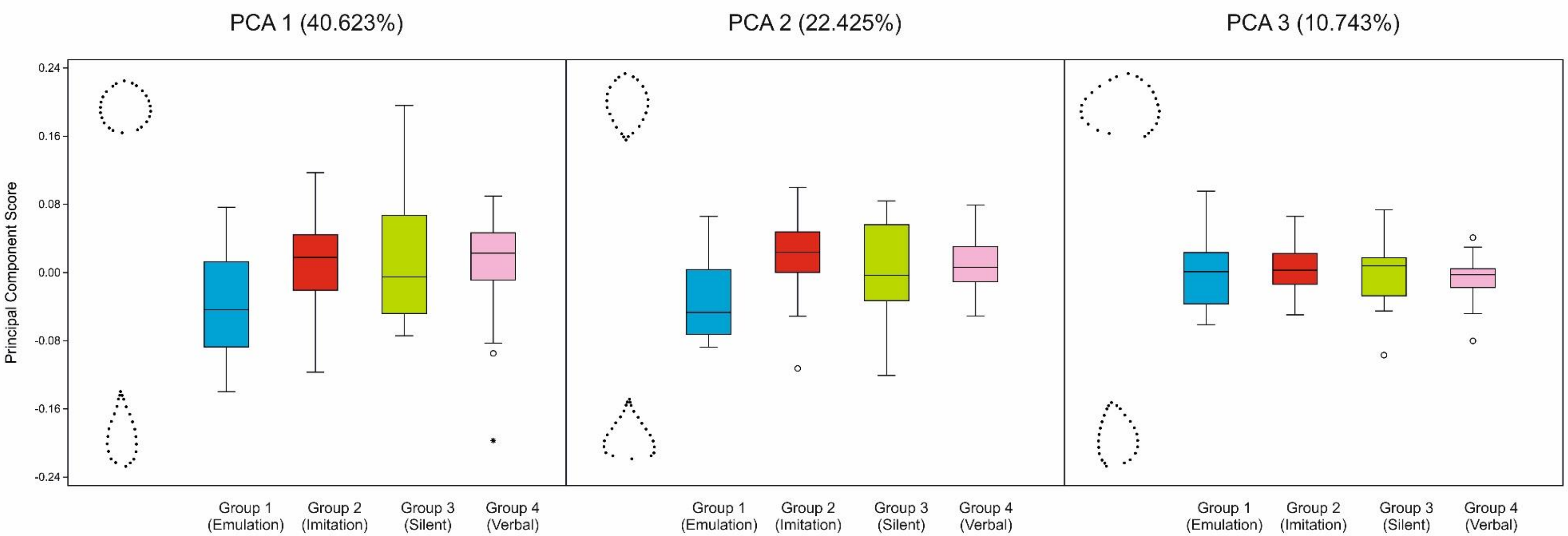
Figure 5.15. A summary of the coefficient of variation values for each variable analysed above

5.4 Geometric Morphometric Framework

A two-dimensional geometric morphometric analysis of the three projections considered in this thesis (planform, lateral and superior views) demonstrates differences in the overall shape of the handaxes produced within each simulated social learning group.

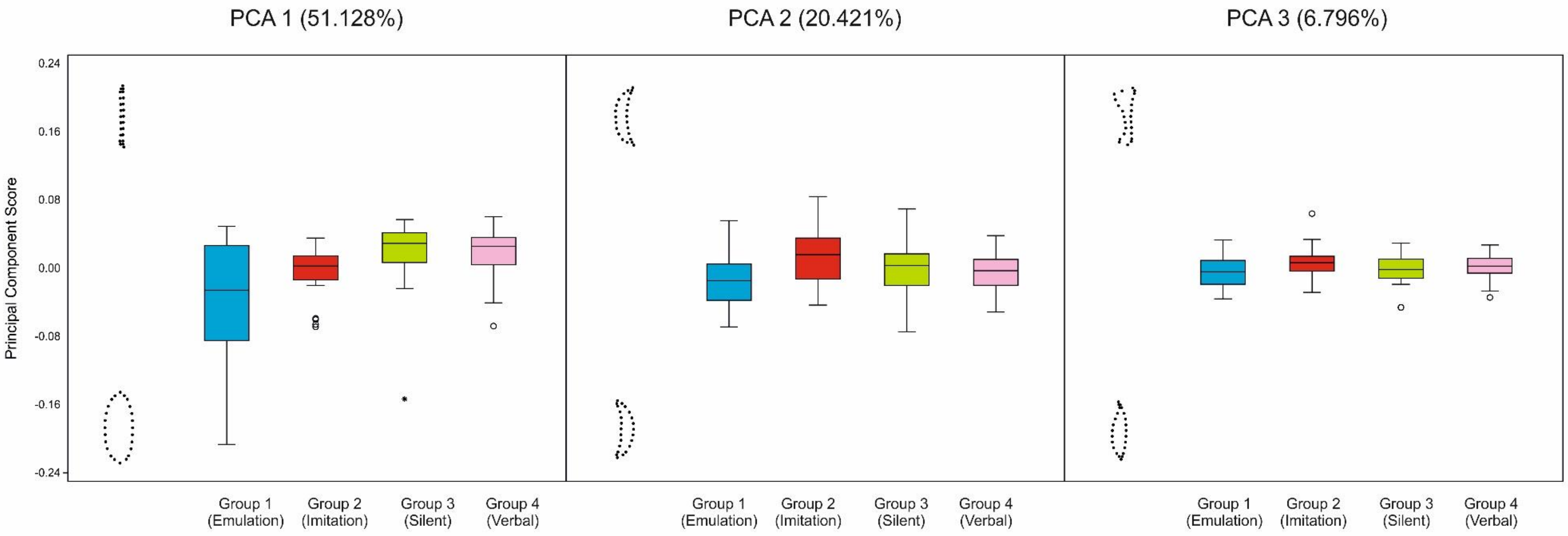
A Principal Component Analysis for planform shape was undertaken, and demonstrated that the first three principal components total just under three-quarters (73.791%) of all displayed shape variation. The first principal component (accounting for 40.623% of all shape variation) describes changes in shape from narrower pointed handaxe shapes to more ovate handaxe types. The second principal component (accounting for 22.425% of all shape variation) describes morphological changes from bottom-heavy to top-heavy planform shapes. The third principal component (accounting for 10.743% of all shape variation) describes changes from a protruding left to protruding right shape. In examining the first three principal components of the handaxes through their planar view (Figure 5.16), the general trend (that was common to the traditional measurements) repeats in that Group 1 (Emulation) has a high, or the highest, range of variation over the three principal components, while Group 4 (Verbal) has the lowest. Group 3 (Silent) seems to not follow this trend and exhibits greater variation than Group 2 (Imitation), and in some cases Group 1 (Emulation). The reasons for Group 3 (Silent)'s irregular pattern is discussed in detail in Chapter Six. A MANOVA of the first ten principal component scores, accounting for 95% total shape variation highlights that Group 1 (Emulation) is statistically significant from all other simulated social learning groups.

A Principal Component Analysis for lateral handaxe shape was also undertaken (Figure 5.17), and demonstrates that the first three principal components total just over three-quarters (78.345%) of all accumulative shape variance. The first principal component accounting for over half (51.128%) of all shape variation describes thicker to thinner lateral handaxe shape. The second principal component accounting for just over one-fifth (20.421%) of all morphological variance extends from lenticular-left to lenticular-right, and the third principal component accounting for 6.796% of all shape variance extends between concave and convex central morphology. With the planar view, Group 1 (Emulation) exhibits a markedly higher range of variation in contrast to the other three simulated social learning groups, which all display similar ranges. In the third principal component, all groups feature a tighter and markedly similar range of variation, however this third component describes a type of morphological variation which is subtler in appearance and as such will not create a large potential for variation. A MANOVA of the first nine principal components, accounting for 95% of total shape variance highlighted that again Group 1 (Emulation) was statistically significant to all other simulated social learning groups, with the two taught groups being the least distinguishable.



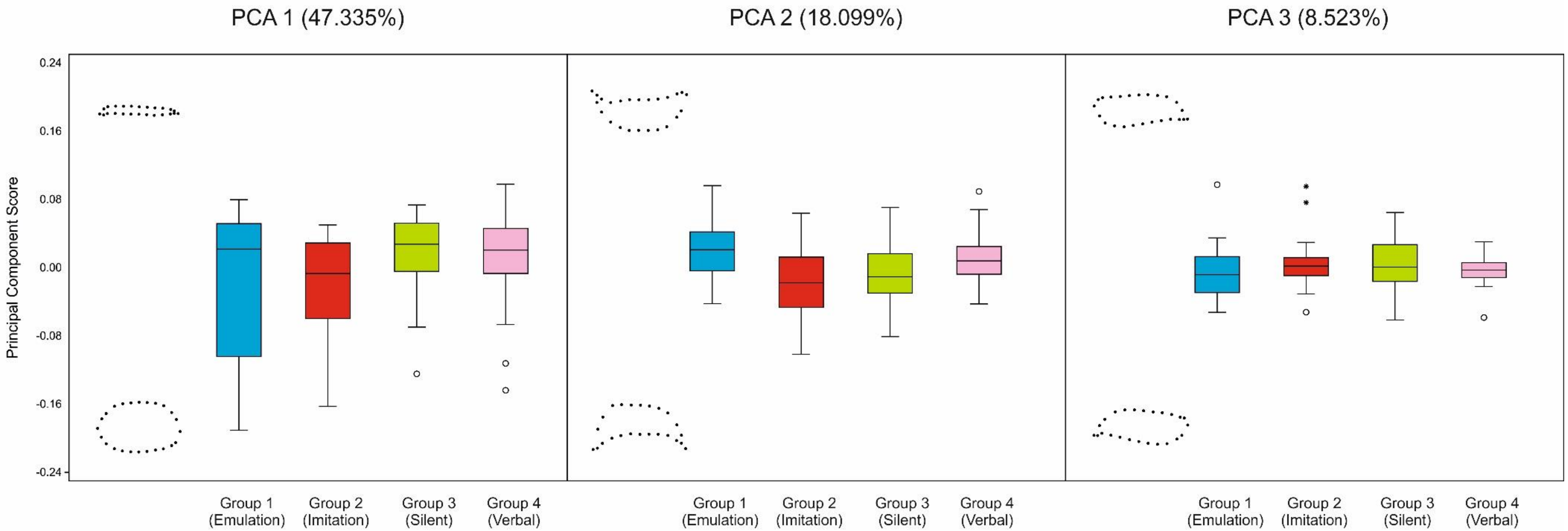
	Group 1 (Emulation)	Group 2 (Imitation)	Group 3 (Silent)	Group 4 (Verbal)
Group 1 (Emulation)		***p* : 0.025**	***p* : 0.045**	***p* : 0.029**
Group 2 (Imitation)	***p* : 0.025**		*p* : 0540	*p* : 0.885
Group 3 (Silent)	***p* : 0.045**	*p* : 0.540		*p* : 0.944
Group 4 (Verbal)	***p* : 0.029**	*p* : 0.733	*p* : 0.944	

Figure 5.16. A box plot of the three principal components for the different simulated social learning groups through their planform projection and a MANOVA of the first ten principal components (accounting for 95% cumulative shape variance)



	Group 1 (Emulation)	Group 2 (Imitation)	Group 3 (Silent)	Group 4 (Verbal)
Group 1 (Emulation)		***p* : 0.012**	***p* : 0.033**	***p* : 0.020**
Group 2 (Imitation)	***p* : 0.012**		*p* : 0.346	*p* : 0.155
Group 3 (Silent)	***p* : 0.033**	*p* : 0.346		*p* : 0.558
Group 4 (Verbal)	***p* : 0.020**	*p* : 0.155	*p* : 0.558	

Figure 5.17. A box plot of the three principal components for the different simulated social learning groups through their lateral projection and a MANOVA of the first nine principal components (accounting for 95% cumulative shape variance)



	Group 1 (Emulation)	Group 2 (Imitation)	Group 3 (Silent)	Group 4 (Verbal)
Group 1 (Emulation)		*p* : 0.115	***p* : 0.029**	*p* : 0.321
Group 2 (Imitation)	*p* : 0.115		*p* : 0.192	***p* : 0.037**
Group 3 (Silent)	***p* : 0.029**	*p* : 0.192		*p* : 0.591
Group 4 (Verbal)	*p* : 0.321	***p* : 0.037**	*p* : 0.591	

Figure 5.18. A box plot of the three principal components for the different simulated social learning groups through their superior projection and a MANOVA of the first ten principal components (accounting for 95% cumulative shape variance)

A Principal Component Analysis for the superior projection (top-down or bottom-up view) was also performed (Figure 5.18), and highlighted that the first three principal components account for just under three-quarters (73.957%) cumulative shape variance. The first principal component accounting for just under half (47.335%) of cumulative shape variance extends from a lenticular to a more concave superior views. In this perspective, tighter clustering can be exhibited within the taught simulated social learning groups, with larger quartile ranges exhibited for Group 1 (Emulation) and Group 2 (Imitation). The second principal component accounting for just under one-fifth (18.099%) of cumulative shape variance extends from convex-left/concave-right to convex-right/concave-left. In contrast to the previous principal component, the overall trend of tighter ranges of variation throughout the four groups cannot be observed. The third principal component accounting for just under a tenth (8.523%) of cumulative shape variance extends from an s-shape projecting from bottom-to-top, and an s-shape projecting from top-to-bottom (see diagram for clarification). In this component, tighter clustering can be observed throughout, with the exception of Group 3 (Silent) which features a range of variance larger than the three other groups. MANOVA of the first ten principal components revealed a more complex picture, in comparison to the previous two statistical analyses, with statistical significance identified between Group 1 (Emulation) and Group 3 (Silent), and between Group 2 (Imitation) and Group 4 (Verbal). While less clear cut than the previous analyses, these still reflect differences between simulated social learning groups which were taught and those that were not intentionally taught. Perhaps most interestingly is the high p value between Group 1 (Emulation) and Group 4 (Verbal), this is discussed in more detail in Chapter Six. A table summarising the statistical tests for all three projections can be observed in Table 5.13.

Projection	Observations
Planform	Group 1 (Emulation) is statistically significant to all other groups
Lateral	Group 1 (Emulation) is statistically significant to all other groups
Superior	Group 1 (Emulation) is statistically significant to Group 3 (Silent); Group 2 (Imitation) is statistically significant to Group 4 (Verbal)

Table 5.13. Statistical observations from analyses of the three different handaxe projections

5.5 Conclusion

This chapter presented both traditional and geometric morphometric approaches to investigate differences in the range of morphological variation between the four simulated social learning groups, and has produced a number of interesting observations concerning their distribution and nature. The variables were chosen as they were anticipated to show differences in the range of morphological variation in both the traditional and geometric morphometric measurements. The analyses here allowed this hypothesis to be assessed visually, including identifying the expected trend of large morphological range in Group 1 (Emulation), and smaller range of morphological variation in Group 4 (Verbal), with the two other groups falling somewhere in between. The interpretations concerning these analyses will be discussed in detail in the next chapter.

Chapter Six.

Discussion

6.1 Introduction

This chapter considers this thesis' research question in light of the results presented in Chapter 5. The research question is:

"Does social learning influence the range of morphological variability of lithic assemblages, and if so can this effect be used to deduce social learning method, theory of mind ability, and therefore language ability amongst Palaeolithic stone tool makers?"

The research question will be answered in the positive, that the results from Chapter 5 show that social learning does indeed influence the range of morphological variability of an experimental lithic assemblage due to differential copying error. With the literature discussed in Chapter 2 and Chapter 3, detection of certain social learning methods in the archaeological record can be used to deduce theory of mind and linguistic ability of the makers of these lithics (with imitation indicating second level intentionality, and teaching indicating third level intentionality).

This chapter will remind the reader of this thesis' purpose and motivation, before summarising the findings and outcomes of the experiment presented in the previous chapter. It then details the importance of the results with respect to its context in the literature alongside the few other studies which consider lithic technology and social learning, and its relevance to the archaeological record and how the themes can be applied. It will also consider at length the cognitive and linguistic interpretations that can be made in terms of broad scale categories of lithic technologies (i.e. Lomekwian, Oldowan, Acheulean and Levallois). This chapter concludes by discussing the further interpretive potential of this thesis while also highlighting limitations to the research.

6.2 Thesis Purpose and Motivation: An Alternative Proxy for Language Evolution

This study was primarily motivated by a need for a new robust material proxy for tracking the origins and evolution of language given the problems associated with the main 'symbolic' method of interpretation (as outlined in Section 1.4). Interpretations of symbolism in the material record leave us with rare, late, and localised conclusions about complex cognition and linguistic ability. The material of focus is also hard to verify empirically as symbolic, and researchers have poorly defined what exact mechanisms tie the material to particular cognitive abilities.

As an alternative, an empirical method of assessing the effects of language ability and complex cognition on material culture is needed. As stone tools are the most omnipresent and abundant artefacts associated with early hominins, they are an ideal material from which to create models. And as cultural objects whose technological knowledge has been socially received, the aspects of social learning and implications concerning theory of mind they bring in tow become a focus for developing material proxies of cognition and language.

This thesis also means to promote experimental frameworks to investigate the past, and create data which cannot be otherwise recovered by excavation or other means. The experimental protocol undertaken in this thesis allowed for the observation of simulated past experiences in a controlled way that allowed for the testing of the hypothesis, as well as the development of new hypotheses. It also developed experimental methodologies for improving the replicability and control of cultural transmission studies involving stone tools with its use of porcelain as a knapping medium.

These theoretical and empirical developments have come together to create a novel study presented in Chapter 5. Its powerful linkages between archaeological material and behavioural inferences contribute to our growing knowledge about Palaeolithic hominins, and offer pathways for future research using these methods.

6.3 Summary of Results

This study has shown that the range of morphological variability within replicated stone tools is affected by the fidelity of the information transfer. Within the experiment, the emulation group consistently produced the largest range of morphological variability, both in the traditional and in the geometric morphometric approaches utilised. The other three simulated social learning groups provided mixed results. They were consistently of a smaller range of variability than the emulation group, but the hypothesised linear progression of smaller and smaller ranges from emulation to teaching did not hold to the expected extent. This could be due to a number of factors, including that the replication task was too simple, and not much else was learned with the aid of teaching. A more complex task might distinguish the imitation and taught groups. Also, the sample size might have been an issue; larger groups creating more handaxes, with longer training times, might result in more differentiation between groups over time. Overall, the Group 2 (Imitation) and Group 4 (Verbal) trend toward smaller ranges of morphological variation, with Group 3 (Silent) not conforming to this trend as often.

While group differences may not have been consistently statistically significant, the emulation group was nearly always the one exhibiting the highest variability, while the verbal group had the lowest range of variability. In the measure of total flake count, Group 2 (Imitation) was statistically different from Group 3 (Silent). However, when the BASM handaxes were removed from the sample, total flake count found Group 3 (Silent) and Group 4 (Verbal) were statistically different. This was not expected, as they had access to the same simulated theory of mind (as they were both taught groups), which involves third level theory of mind (although their teaching styles were different). However, the silently taught group could have produced less flakes due to poor silent teaching, or from being in an awkward situation where miming was being used for the teaching. In the superior projection, Group 4 (Verbal) was found to be statistically different from Group 2 (Imitation), and Group 3 (Silent) was found to be statistically different from Group 1 (Emulation). For the reasons stated above, these differences between the imitation group and the taught groups might be further accentuated with changes to the experimental design that introduce more complex variables to be replicated.

Also, due to time constraints, the analysis did not include observation of how the groups differed when only later replication attempts were considered, such that progressive attempts might have had increasing faithfulness to the model by varying degrees. Therefore there might possibly be a different level of similarity between groups when only considering later attempts. A follow-up analysis that considered this possibility could prove interesting, and it is hypothesised that the more practise the groups have, the more reduced the tool variability might be (and perhaps to different extents), as they will converge toward the model more as they learn to create more faithful copies.

Results of this thesis demonstrated the complementary nature of using both single variable traditional morphometrics and methods of shape analysis with geometric morphometrics in order to assess the range of variability. Relationships between groups showed similar results, and highlighted the variability of different features of the material.

In returning to the research question, the results from Chapter 5 showed that the morphological variability was greatest in the low fidelity simulated social learning environment of Group 1 (Emulation), and consistently the lowest variability in the high fidelity simulated social learning environment of Group 4 (Verbal). This demonstrates social learning's effect on morphological variability: that the fidelity of the knowledge transfer is realised in the morphological variability of the replicated tools, and that standardisation is a feature of high fidelity social learning. The replicative nature of this study offers many avenues for continued research (discussed further in Chapter 7), and the hypothesis has great potential for further support resulting from future experiments and archaeological applications.

6.4 Contextualising the Research: A Consideration of the Wider Literature

The results from other studies examining social learning in the context of modern lithic experiments were discussed in Section 3.3.2. When taken within the context of the results from this study, we find consistency. In Schillinger et al. (2015) and the experiment within this thesis, the lowest fidelity learning situation (emulation) had the least amount of standardisation. In Morgan et al. (2015) and this thesis' experiment, the verbally taught methods were found to improve the transmission of the task. In Morgan et al.(2015)'s study, this meant successful transmission of knapping a flake, while in the present study the verbal group featured more standardisation, as the features of the knapped handaxes were more faithfully replicated and were therefore more alike each other.

Along with the present study, Ohnuma et al. (1997) and Putt et al. (2014) compared silently taught groups to verbally taught groups. Both Ohnuma et al. and Putt et al. found little difference in the acquisition of knapping techniques. However, because the starting material was not of a standardised form, the final tools could not be reliably assessed for shape comparison, as starting shape and stone quality would all play a part in the final form. Putt et al. (2014) did note that they received an impression, however, that the verbal instruction did not have a significant effect on the shape of the tool participants created. These findings differ from this thesis' experiment, where the range of variation between the silently taught group and the verbally taught group were often somewhat different (though rarely significantly so). It is understood that people can learn to knap with silent instruction, and this is what Ohnuma et al. (1997) and Putt et al. (2014)'s studies demonstrate. However, these experiments do not stand as experiments that show how knapping technical information can be transmitted by and to non-linguistic individuals. Silent transmission is not a viable simulation for a non-linguistic brain. The hypothesis of this thesis' study was that because of the similar theory of mind at work in gaining knowledge about how the handaxes were made, there would not be significant differences between the range of variability in Group 3 (Silent) and Group 4 (Verbal). As the unconventional situation of learning by miming and gesture, as well as the lack of experience and perhaps comfort in asking questions by miming and gesture (and the lack of cultural conventions for doing so), the participants in Group 3 had an impoverished exchange of intentional communication in comparison to Group 4. This imbalance is what could have caused the disparity in the range of morphological variability between groups in certain variables. Both groups, however, did experience more clustering and a lower range of variation than Group 1 (Emulation), and less so from Group 2 (Imitation).

6.5 Interpretive Potential of this Thesis

The fundamental importance of the work of this thesis is the contribution to the understanding of the nature of lithic variability and standardisation, and its relationship to social teaching mechanisms. It offers new insights and new gives new meaning to what lithic variability means within archaeological analyses of hominin behaviour. This thesis also highlights the importance of a rigorous experimental framework within studies of language origins. Interdisciplinary lines of evidence must converge and synthesize to ensure ideas of language's origins are compatible with evolutionary theory (Kinsella 2009) and theories of neurological and psychological development, at least to understand the cognitive co-evolutionary mechanisms that are possibly at play. Because language is ephemeral, we cannot excavate it in the same way as we do lithics and other material culture. Experimental protocols therefore offer us the best way to gather data through simulating past events and controlling the variables therein. Porcelain was used as a raw material source in this thesis' experiment for precisely this reason, to introduce a scientific control that also aided in replicability of results, and a material that offered standardisation which allowed the morphometric analysis of results.

Finally, this thesis' showed how both traditional and geometric morphometric approaches to lithic variation can be used to show a complementary picture of lithic variation. The morphometric analysis conducted in this thesis highlights in handaxes a planar view is not the only perspective which aids in our understanding of lithic variability. Indeed it was the superior (top-down) perspective that found a statistical difference between Group 2 (Imitation) and Group 4 (Verbal), a distinction not observed in the planar or lateral perspectives of the handaxes. As archaeologists we experience handaxes as objects flat on a table, their planar shape as our stereotyped idea of their regular form. But to the people who made and used these large cutting tools, which were held and manipulated in the hand, these tools would have been experienced in a much more three dimensional way, as very active objects to be engaged wtih. The edge angle of the cutting edge of large cutting tools for example, is best seen through its superior projection. The analysis of the handaxes in these three different perspectives has hopefully supported why the planar view should not be the only perspective analysed when taking tool shape into account.

6.6 Limitations of the Study

It is necessary to highlight any shortcomings of this thesis and the experiment so that conclusions may not be over-interpreted, and issues are identified that might have biased results. One issue is that this thesis is relying on studies with modern participants to make inferences about extinct hominin species. However, the author argues that the nature of theory of mind and language itself suggests that these two would not exist separately under any condition, as one emerges almost as a consequence of the other. Language and theory of mind are both social functionaries that involve the meeting of minds in dyadic relationships, and the operations of one cannot be conducted without the operations of the other (which are often one and the same), and this is illustrated well in the common precondition of joint attention, which enable the growth of both abilities.

Another potential weakness of this thesis could be said to be its inferential nature. As we are using a chain of inference to connect lithic replication to social learning method, to theory of mind, to linguistic ability, then the theory falls apart if any of the links are broken; the theory will only be as strong as its weakest inferential link. However, it offers a degree of falsifiability which is necessary and welcome when considered next to the symbolic proxy for cognitive ability, or other studies which are purely theoretical and have no direct applicability to the archaeological record. Its falsifiability is its strength as a scientific endeavour.

The discipline of theory of mind research in itself has theoretical issues, in regards to how we understand its mechanisms and development. There will surely be changes as new research develops. Theory of mind had a surge of research in the 1990s, when modular and cognitivist views of the brain and mind were popular. As a result, theory of mind and much of its research still reflects these views. However, theory of mind is not inherently cognitivist (Stade and Gamble, *accepted*) and is fully workable with a materialist, embodied mind approach taken here. It is hypothesised that the step-like nature of levels of intentionality in theory of mind will begin to be explored further, and that there will be more research into how, for example, desire understanding might be acquired earlier than belief states (Liu et al. 2009, Wellman et al. 2005). This more nuanced view of theory of mind will lend itself to new interpretations to its social learning and linguistic links.

A commonly cited problem, also of relevance to this thesis, regards issues of sample size. The number of participants and handaxes were constrained by time and funding, though not as constrained as they would be if porcelain was not the material used, and a more expensive and difficult to prepare alternative of flint was used. A future study with more participants creating more handaxes would be greatly beneficial to see if the learning plateaus for example, or if a larger sample size gives more or less significant results and ranges of variability (discussed further in Chapter 7).

As this study involved the training of participants, the abilities of the trainer are an influencing variable. A teacher who had strong acting abilities might have conveyed information in the silent teaching group more successfully for example. Also, the author is not an advanced knapper, but has only been knapping for 4 years. A more advanced knapper might themselves have a more idiosyncratic knapping style that would be better transmitted to the participants in the study, and be more consistently presented.

As for limitations of the analyses, one is that the geometric morphometric methodology did not offer a clear way to quantify the range of variability for comparative analysis, since the analyses are all relative to each other. A better method of quantifying the range of morphometric variation is desired, and this is the logical next step for this study in terms of making it more applicable to the archaeological record"Does social learning influence the range of morphological variability of lithic assemblages, and if so can this effect be used to deduce social learning method, theory of mind ability, and therefore language ability amongst Palaeolithic stone tool makers?". There was also no consideration of attempt number in the analyses due to time constraints, or isolating later attempt numbers to see if there was a temporal relationship, and these are opportunities to explore in future analyses, also discussed in Chapter 7.

6.5 Integrating the Research Framework with the Archaeological Evidence

Finally, we will discuss how the information from this thesis can be applied in a useful way to the archaeological record, and the interpretations arising from the results in Chapter 5 that predict broad scale cognitive and linguistic ability within certain lithic industries throughout the Palaeolithic, in an effort to characterise the potential teaching-learning environments used within them.

The goal of this thesis is to develop a proxy that can ultimately be used against the archaeological record to make interpretations about cognitive and linguistic ability from archaeological materials. In order to do this, the understandings of how social learning methods affect standardisation need to be understood and then applied. This study looked at the replication of bifacially shaped porcelain artefacts, and it is parsimonious to assume that these variables are at work in other mediums (such as flint and other knappable materials) and on other types of tools. These results are pertinent to more than just bifacially worked stone (or just stone tools in particular), and in replicated material more generally, including organic materials. The potential exists as well for an analysis of use-wear patterns, as they are the result of the physical replicated processes being applied to the material.

As the results from the experiment show, morphological variability decreases with higher fidelity forms of social learning. We gain a better understanding of another of the variables that affect the morphology of Palaeolithic assemblages. We can now use this information to identify those assemblages that could only have been produced by high fidelity social learning. A number of variables affect how similar or different replicated artefacts will be to each other, and most variables will introduce increased variability (such as raw material and function). High fidelity information transfer, however, is one of the few variables that will actually allow for an increased similarity between replicated objects, because of how similar physical gestures will cause similar physical consequences. Replication of an object includes copying of the motor sequences that produced the initial object, and so the more faithful the sequence to the original, the more similar the replicated object will be. In stone tool technology, then, high fidelity cultural transmission of an object through imitation or teaching will have the potential for high levels of similarity of the replicated object. Within cultural materials, high levels of similarity will therefore indicate high fidelity forms of cultural learning have taken place. And this means that interpretations about the social learning method and its coordinating theory of mind and linguistic abilities can also be interpreted from the level of standardisation.

The applications of this work are therefore to interpret teaching and learning in the archaeological record so that 1) broad industry-level assessments of cultural transmission ability can be assessed, and the cognitive implications they necessitate, but also 2) conducting assemblage-level assessments of tool variability, and understanding the manifold variables that impact shape and form of products of knapping sequences, so that an individual horizon or site level assessment of the replication in a community or close communities can take place.

Any major distinguishing units of Palaeolithic lithic technologies are necessarily reductive, and to an extent, reflect our own biases as modern archaeologists in fitting technological evolution into discrete categories. These categories are based on features we deem the most salient and meaningful to our analyses regarding their temporal, spatial, and morphological aspects. As Clark (2009) notes (referring to terms such as Aurignacian, Mousterian etc.), they are 'accidents of history', and analytical units devised to solve problems of chronology before absolute dating techniques were available. *Lomekwian*, *Oldowan*, *Acheulean*, and *Levallois* are four such lithic industries that will be discussed here. The categories are being used as heuristics of commonly accepted technological methods because of familiarity with the readership. Their associated technological features are the subject of how they might have been successfully culturally transmitted; they are not, however, meant to reflect absolute paradigm shifts, which is a matter beyond the remit of this thesis. In this way, though, they offer us convenient categories, and do not strive to divide millions of years of lithic variability into teleological or spatio-temporally discrete categories. Based on the hypothesis proposed by this thesis, the social learning implied by the behaviours required for the successful transmission of these knapping techniques will now be explored and interpreted in terms of their required theory of mind and language.

6.5.1 Lomekwian

In 2015, information about a newly discovered lithic tradition was published, which pushed back the starting age of the Palaeolithic period. The site predated the previously oldest known stone tools at Gona, Ethiopia (Semaw, 2000) by 700,000 years. Lomekwi 3, on the shore of Lake Turkana in Kenya (Harmand et al. 2015), has been dated to 3.3 million years, and the term for the tool industry found there, distinctly different from later Oldowan sites, has been termed Lomekwian.

Before the Lomekwian tools' discovery, researchers had long speculated about the undiscovered existence of previous stone tool industries (e.g. Panger et al. 2002), because of Oldowan tools' possible associations with a number of potential different hominin species. Its apparent sophistication made researchers suspect an earlier, simpler technology (Delagnes and Roche 2005). Hayden (2015) predicted that an earlier, bipolar-based industry might have existed which predated Oldowan technologies, as core and flake tools similar to Oldowan tools are ethnographically known as useful woodworking tools, and bipolar technique similar to Lomekwian tools are deemed the most expedient way to create quick, sharp edges for butchery.

Hayden also suggests that such a technology might provide "a key link between primate nut-cracking technologies and the emergence of more sophisticated lithic technologies leading to the Oldowan" (Hayden 2015, p. 1). Experimental replication of the tools at Lomekwi 3 suggest that they were made by passive hammer and/or bipolar technique, and Harmand et al. (2015) note that these are techniques rarely found associated with Oldowan sites; the flakes and cores are mostly large and heavy, compared with Oldowan artefacts, or accidental flakes made by Chimpanzees from nutcracking. A number of passive elements (anvils) and active elements (hammers) are included in the assemblage, and battering has been noted on some of the flakes on their dorsal side, implying percussive activities.

The links to percussive technology are certainly interesting, and it can be seen why Hayden (2015) would suggest that a bipolar early stone tool industry would provide a more continuous link to stone technologies used by other primate species. Because of these similarities in the technology, we can also make inferences about the social learning requirements of such a technology. If chimpanzees are able to sustain and transmit a bipolar lithic technology, then Lomekwian tools too do not offer a strong case for the requirement of anything more than emulative social learning behaviours for its successful persistence in the culture. Chimpanzee nut cracking technology does not include intentional bipolar fracturing of stone, but replace the nut with another stone, and the technique is virtually the same. A passive hammer technique to detach sharp stone flakes requires the same motion, but with the absence of a nut/stone. The mechanisms involved are of quite a similar kind, and it is feasible that young hominins could acquire the technologies in much the same way young chimpanzees acquire the ability to manipulate the passive and active elements of stone in nut cracking chimpanzee communities.

In a hominin community that did not utilise imitative behaviour in order to transfer knowledge about stone knapping, it is likely that this is because the ability was not available, rather than just not utilised for this specific purpose. Imitative learning would be eminently useful in the transfer of this knowledge, and if it could be applied, it most likely would be. Therefore assessing what looks to be an emulative learning process such as that simulated in Group 1 (Emulation) and interpreting the absence of an imitative ability does not seem illogical (as opposed to a usual 'absence of evidence' scenario). What this means for theory of mind, then, and ultimately language, is that Lomekwian hominins were not sharing mental worlds through a theory of mind; and within that, they were not communicating those mental worlds with linguistic labels for things.

6.5.2 Oldowan

For decades, and up until only recently (Harmand et al. 2015), the Oldowan was the oldest known knapped stone tool industry, taking its name from its place of discovery, Olduvai Gorge in Tanzania. Sites range in age from 2.6mya (Semaw 2000) to around 1.6mya (Wynn et al. 2011). Its typifying technological attributes have been found to be difficult to characterise, and Braun and Hovers note that the industry is "marked by its diversity" (Braun and Hovers 2009, p. 1). There is even disagreement over whether the Oldowan constitutes a single industry or would be better described with different spatial, temporal, or technological distinctions (Stout et al. 2010). The Oldowan differs from the Lomekwian mainly in that its technology is characterised as having been produced free-hand; a core held in one hand is struck with a hammer in the other to produce a flake with a sharp cutting edge. These flakes can appear in a combination of parallel or alternate flaking techniques (McNabb 2007). Small flakes are characteristic (Barsky 2009), and tend to be the product of the modification of a number of different core form shapes (Leakey 1971) including spheroids.

It is this core management that has led researchers such as Braun and Hovers (2009, p. 3) to assert that, "The very early lithic assemblages in the growing Oldowan database already followed at least rudimentary principles of geometric and spatial planning of lithic reduction.". Following this, some have speculated that Oldowan technology suggests advanced cognitive attributes such as teaching and language (Hovers 2012, Bickerton 2009), although Wynn et al. (2011, Wynn and McGrew 1989) suggest that imitation, or mechanisms found in chimpanzee communities, might be enough to transmit the behaviours. As wild chimpanzee communities appear not to use imitative behaviours, it could be assumed that Wynn et al. then consider emulative behaviours to be sufficient: "At most one can argue that the Oldowan pushed the limits of ape grade adaptations; it did not exceed them." (Wynn and McGrew 1989, p. 394) However, Oldowan tools do not exhibit features that suggest attention to the overall shape to the point such as is typically found in the Acheulean, as concluded by Toth (1985, p. 101): "many Oldowan core forms ("core-tools") are, probably simple by-products of flake manufacture rather than representations of stylistic norms."

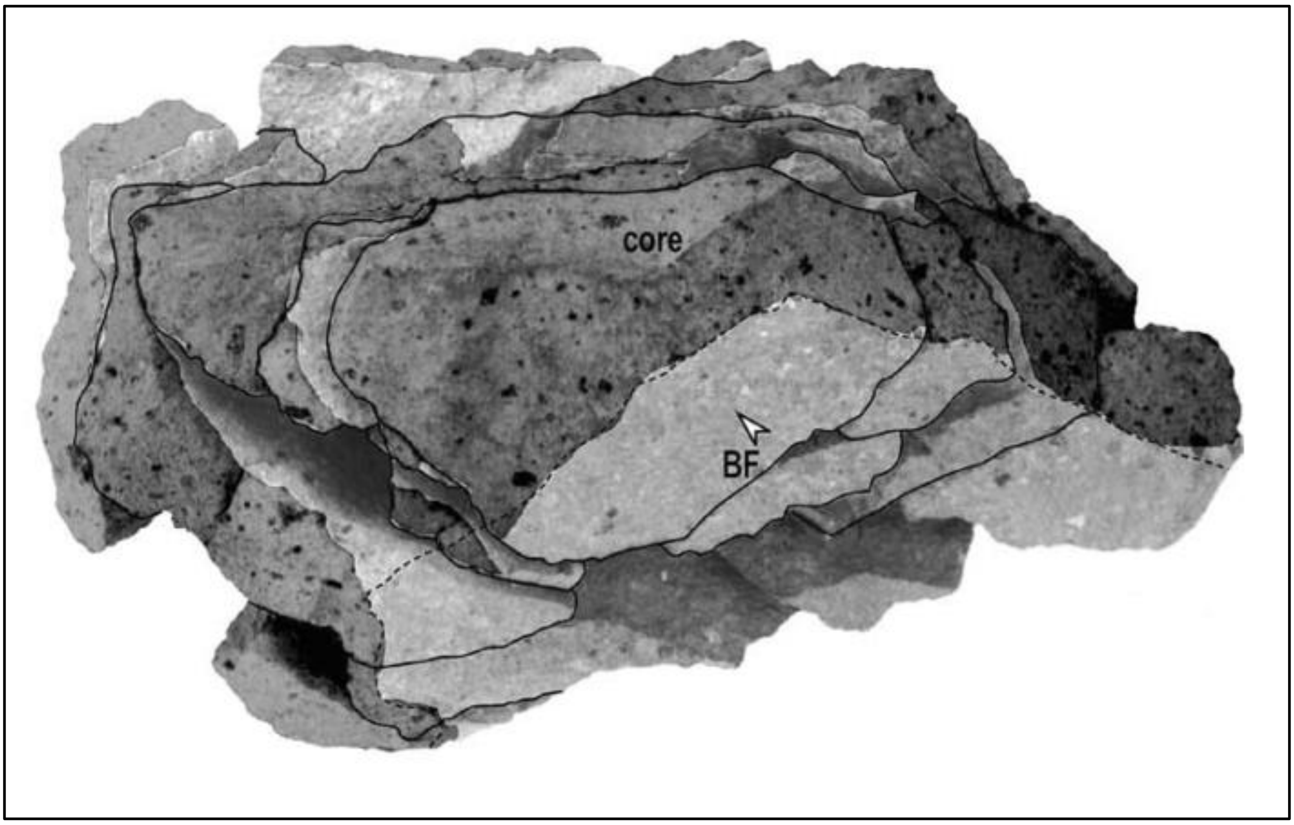


Figure 6.1. A refit group from Lokalalei 2C made up of 38 items knapped on a basalt cobble. From Delagnes and Roche (2005).

Some Oldowan sites have produced impressive assemblages that have allowed archaeologists to refit knapping sequences that show more than 50 removals from a single core (Delagnes and Roche 2005), such as those at Lokalalei in Kenya. Lokalalei 2C is a famous example of this early sophisticated usage of raw material to produce abundant sharp edged flakes. At this site late Pliocene hominins brought in more than 100 locally sourced cobbles, where flakes have been refitted showing an average of 18 flakes per core. Impact damage shows stabilised motor control and precision, and researchers interpret the knapping behaviour as showing foresight and planning. They describe the flaking procedure as being guided by 'technical rules', which evokes behavioural similarity of the type that might have come about with imitation; so while the tool forms themselves are not idiosyncratic, the techniques (such as alternate flaking where the scar from the previous flake becomes the platform for the next) might have been.

The poor quality raw material, but expert way in that it was exploited for a maximum number of flakes, signals that these hominin knappers were not novices, and quite possibly using cultural knowledge developed through a cumulative 'ratchet effect', rather than skills learned over years of individual experimentation and personal innovation (or reinvention of the wheel). And as the knappers at this site were plural (18 hammerstones recovered signals the likelihood of more than just a few knappers, assuming contemporaneity of the assemblage), means that this stereotyped set of technical rules could be of the kind only shared by a community that transmits knowledge in a more sophisticated way, such as with imitation, which allows that ratchet effect such that a 'set of technical rules' can not only form, but be acquired and shared within a community. As imitation is built upon joint attention and the understanding of intention, the hominin group at Lokalalei could be characterised as hominins with the symbolic capacity for creating and understanding shared linguistic labels for things.

Whether or not some Oldowan technologies represent a clear case of techniques that require transmission of a different sort than that used to produce the Lomekwian assemblage, or the learning of nutcracking in chimpanzee communities, is less clear than the next technical example we will turn to, which offers a stable, recognizable form that allows for the measurement of shape variability.

6.5.3 Acheulean

Many see the emergence of core tools as a paradigmatic shift between simple flaking methods not necessarily meant to produce core forms, to the emergence of serial knapping events with a strategic imposition of form and the resultant core tool as its goal. The Acheulean is characterised by the presence of large bifacial cutting tools, including the classic 'handaxe', but also cleavers and picks. This period of bifacial core-dominated assemblages begins at around 1.7mya (Asfaw et al. 1992), and continues to about 200,000 years ago. While the technology is often characterised as a period of 'stasis' (Hopkinson et al. 2013), the later Acheulean does exhibit some handaxes of standardised, fine quality, created with thinning flakes, soft hammer percussion and prepared platforms such as at Boxgrove (Stout et al. 2014). However, the Acheulean being marked by a progression to more standardised handaxes over time, or more symmetrical, while sometimes suggested, is not a simple consensus view (Cole 2015, Iovita et al. 2017).

Regarding Acheulean technology in a social context, Wynn (1995 p. 20) writes,

"Making a handaxe requires another step up in complexity, for actor A must not only conceive of what actor B sees, he must also be able to conceive of what actor B believes to be an acceptable final shape. Even if actor A watched B produce scores of examples, this would not in itself result in A producing the same range of standard shapes, if A could only conceive of what B saw. A must somehow come to know what B understands to be appropriate. He must know that such a belief exists, and just what that belief is. This is much more difficult than simply constructing a visual perspective, because it requires constructing the content of another's mind. This is intersubjectivity. But is it possible to transfer such a complex 'what' without recourse to symbolic communication?"

Without explicitly using the words theory of mind, social learning, cultural transmission, or copy error, the author is alluding to these processes and the way in which the learner must perceive the actions of another as intentional in order to see them as functional and to be replicated, and the passing of information from one to the other as not a direct copy, introducing variation. Then the author alludes to symbolic communication, language. The link between social learning, theory of mind, and language, has seeped into this thought exercise about the transmission of knowledge specifically regarding handaxes, which are the first lithic tools that researchers can confidently ascribe a stereotyped 'form' to.

Lycett et al. (2016a, p. 31) considered that, "Imitative learning of specific manufacturing details would also explain why metric studies of handaxe form often display statistically patterned variation between assemblages of handaxes from different regions or sites." As this thesis is measuring exactly that (statistically patterned variation under a controlled experiment which creates different assemblages), provided here is empirical support for Lycett et al.'s consideration.

Chazan (2012) considers when in human evolution we can first identify teaching. He believes that handaxes appear to show signs that hominins approached the making of the tool with a strategy in mind, and that *Homo erectus* would have been learning 'concepts' of handaxes learned through the teaching by others. Handaxes tend to be the earliest tools that researchers are willing to ascribe these 'teaching' behaviours to. McPherron (2000) questions how much of a learned component is involved in the making of handaxes, and how much can be deduced from their form. As so many other variables affect their form, including raw material, variation could mostly be a function of these other factors.

The stability of form in the Acheulean has motivated the conclusion from a number of researchers that the transmission of its technical knowledge was a product of imitative learning, and this was what allowed for its stasis of form (Shipton et al. 2009). However, Corbey et al. (2016) questions whether we can be sure that handaxes are cultural objects, in that they have been learned by social learning processes, at all. As Chazan (2012, p. 198) writes, "Handaxes are clearly something the human, or even hominin, mind "does well" but at the same time it is also clear that making handaxes is not a natural part of ontogeny." Corbey et al.'s paper presents the most robust argument to date that handaxes might not constitute a clear example of cultural transmission (or at least that we have not yet established the case well enough that they are, and we instead take the point for granted). They consider this general agreement among researchers, that handaxes are cultural, and not genetically influenced. However, they note that Richerson and Boyd (2005) point out that the stability in form of handaxes could possibly support the opposite of being the product of cultural knowledge. Corbey et al. suggest that copying error would introduce much more variation than that exhibited in handaxe variation over time. They say that an experimental study (Kempe et al. 2012) showed lower copy error than expected from a large sample of archaeological handaxes. However, a conclusion could be that the stability in form itself could be the result of a step-change in the complexity of the socially produced tools, and that the later proliferation and diversification in forms is a product of a stronger ratchet effect allowed by teaching. A low diversification caused by low ratchet effect could imply imitative learning, not teaching. We have no ethnographic or non-human examples of cultural change motivated simply by imitative learning, and so we cannot be sure that this archaeological signature is not just the result of an imitative hominin, and a persistent culturally transmitted tool type. Social learning of this kind could create the illusion of a genetic component due to its stability.

Social learning is not of one kind. Different cognitive mechanisms affect the variability of the transmission, as we have seen from the experimental results shown in this thesis. If Acheulean technologies were governed by imitative learning and cultural transmission, we would expect to see stability in form allowed by the transfer of technical knowledge, and innovation derived from copy error, but also innovation and cumulative cultural evolution allowed by a ratchet effect (although it is unclear how much cumulative cultural innovation is allowed for by imitative ability, and this would surely be an interesting area for further research). For instance, it is not known whether this is what would predict or contribute to the ever discussed 'stasis' during the Acheulean. Furthermore, stasis could simply be relative to what we see compared to the rapid cumulative cultural change that follows during the Middle Palaeolithic.

From the argument built upon the evidence from this thesis, Acheulean hominins were likely using imitative behaviours similar to that simulated in Group 2 (Imitation) to socially transmit cultural and technical knowledge. This interpretation is very impactful on the other behaviours we can then deduce. As imitative beings, these hominins would have the ability not only to engage in joint attention, but to operate in a world where they appreciate what others are thinking and feeling (or at least what is suspected), and what knowledge and beliefs others hold. This opens up a critical shared mental landscape, where material items (including people) can be ascribed common meaning within a group (such as linguistic labels for things). Minds are no longer alone. A shared symbolism is created, as appreciations for others' mental experiences can be catalogued and rated by members of the group. Associated labels emerge when common referents are shared, and understood to be shared within dyadic relationships; imitative hominins would have words.

6.5.4 Levallois

The most often cited technology that researchers have suggested might implicate a form of teaching is Levallois, and presumably other types of prepared core technique (for an interesting discussion see Schlanger 1996). Because of the shaping that is applied to a core to detach one or a number of predetermined flakes, Levallois has been seen as implicating certain planning abilities and the necessity of verbal communication for its acquisition. It therefore offers the chance to examine, more explicitly, questions of cultural and intentionally transmitted ideas. Levallois and other prepared core techniques work on the basis of 'predetermination' of a core that has been shaped such that there is a somewhat domed area of raw material that the striking energy will remove. This can be for one or a number of recurrent flakes as the domed shape of the core allows.

Lycett et al. (2016b, p. 20) recently reviewed work undertaken on social learning associated with Levallois technology. They suggest that, *"In the case of Levallois, a requirement to effectively learn both aspects of core form and the successful reproduction of extended and strategically organized knapping sequences, would again indicate that – minimally – combinations of emulation (results-based) and imitation (action based) social learning were required."*

If this is so, then the Middle Palaeolithic would be minimally characterised by a simple referential and symbolically mediated cognitive life, with linguistic labels for things and the shared mental landscapes that necessitates. But the techniques utilised in preparing a strategic platform to detach a preconceived flake, if not only technological but functionally necessary to be explicitly intentionally transmitted, implies a suite of other much more complex cognitive attributes that make Middle Palaeolithic toolmakers likely to have a much more rich cognitive and linguistic life of the sort that implies the possession of third level theory of mind. This does not mean that Levallois technique had to have been transmitted verbally (or through other linguistic mediums such as sign), but that it probably was the case, because the cognitive ability to intentionally transmit information is built upon the existence of a linguistic community wherein intentional conferring of knowledge could operate through complex theory of mind.

It is also possible that these features existed prior to the Middle Palaeolithic; complex stone tool techniques with functionally opaque behaviours such as edge preparation for thinning flakes would minimally require over-imitation, to explain the propagation of an opaque feature without the understanding of its function; but without the intentional conferring of techniques such as these, it is possible that they would not transfer successfully often enough to persist in the record at Lower Palaeolithic sites such as Boxgrove, where there are a multitude of thinned handaxes with low morphological and technological variability (Stout et al. 2014). It is possible to reasonably extend the argument for intentional teaching and third level theory of mind of the kind simulated in Group 4 (Verbal) back to the late Acheulean, making the origins of grammatical language (which for many linguists who see grammar as the all-important Rubicon for a communication system to be designated as language) a very deep origin indeed, well before conventional estimates of 70-100ka (e.g. Hauser et al. 2014).

As this thesis is concerned with theory of mind levels of intentionality up to level 3, where we have established this at a likely Middle Palaeolithic or late Lower Palaeolithic phenomena, we will not continue looking for further technological complexities that may require yet further complex theory of mind abilities to appreciate, although social learning abilities that characteristically require level 4 intentionality are not known to the author. The implications of traditionally discussed complex material culture such as hafted implements and multicomponent transformational materials such as adhesives would be an interesting topic to explore in future in terms of what they tell us about the transmission of cumulative cultural information.

Another point to note is in regarding standardisation through time. Handaxes, for example, do not appear to show increased symmetry or refinement over time (McNabb and Cole 2015). The model presented in this thesis does not predict that lithic assemblages would experience increased morephological standardisation through time, as a number of variables impact artefact variability and many could be at play at once, however complex the social learning involved. However, what this model does predict is the appearance of occasional assemblages with levels of standardisation higher than that which appeared before, where there are special cases when variables that affect variability (including social learning) are not increasing that variability. Complex social learning is, rather, necessary but not sufficient for a Palaeolithic community to produce many like tools; therefore, this model only predicts occasional appearances of highly standardised material, and not a universal trend.

6.5.5 Conclusion

Taking into consideration the technological behaviour of different lithic industries, Lomekwian technologies do not present a behaviour that is of a different kind than could be supported by an emulative social learning system similar to chimpanzee nutcracking technologies. Oldowan core and flake technologies present a more stereotyped behaviour, though it has been noted that the industry is marked by behavioural variation. Some Oldowan sites, such as at Lokalalei 2C, present complex knapping behaviours and expert control of material despite their deep antiquity. As the intention seems to be more for flake production than to shape a core, the analysis of the resulting material assemblages does not offer the same opportunity for analysis that core tools such as handaxes do. However, it is feasible that some Oldowan knapping behaviours necessitated imitative behaviours to learn to exploit cores in such an effective way. Acheulean handaxes provide a more confident case for the transmission of tool production knowledge that requires more complex learning mechanisms like imitation, and so it is likely that Acheulean hominins possessed abilities such as theory of mind, and symbolic abilities like word reference which scaffolded these imitative abilities. Acheulean technologies are often characterised as static, but the persistence of form might itself be a consequence of imitative learning that allowed for a persistence of technology. Not all handaxe technologies are the same, however, and in the later Lower Palaeolithic, certain (but not all) examples of handaxe technology show features such as thinning flakes, which require edge preparation (for example at Boxgrove). They also show a low level of morphological variability. Examples such as this could indicate techniques that had been transmitted in a more complex way than solely imitative behaviours, with intentional conferring of knowledge. Prepared core technique such as that of Levallois is a technological behaviour often indicated by researchers as being of complexity that might require verbal delivery of knowledge, or otherwise taught behaviour. The skill and planning involved has been interpreted as alluding to complex cognitive abilities by various researchers. These opaque techniques (those not easily understood by visual cues alone) of planned flake removal, and the preparation for future removal, are not as explicit or connected in immediacy to their function as other techniques; their imitation and persistence in a community may therefore depend on teaching to acquire the understanding of a perceived function, which might remain unknown without the explicit intentional conferring of that specific knowledge.

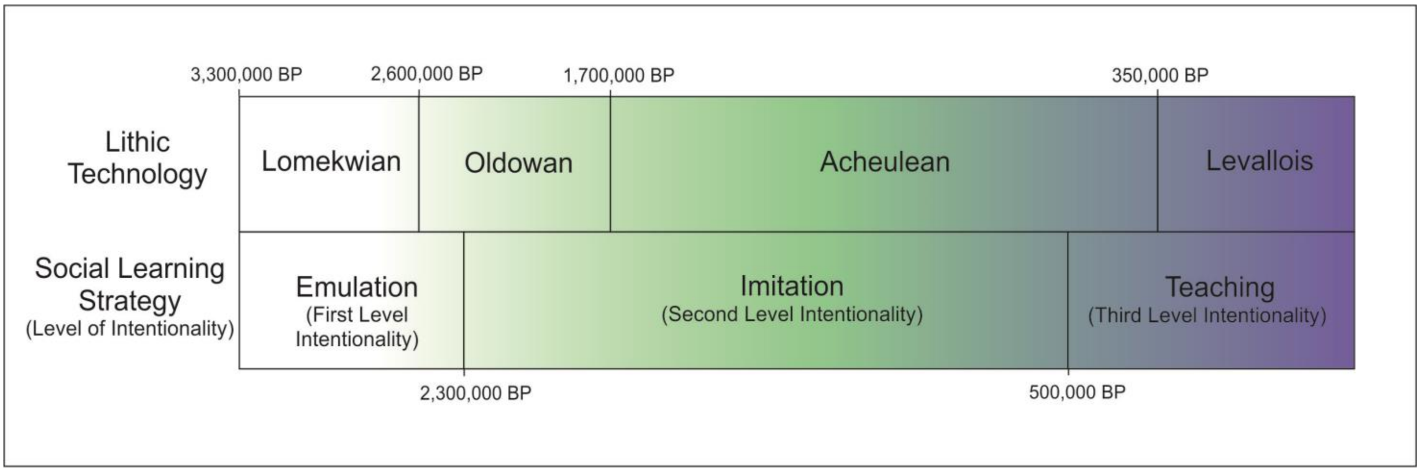


Figure 6.2 A broad schematic showing the proposed correlations between lithic industries and correlated social learning abilities.

This would require theory of mind on the part of the teacher in order for the learner to gain that knowledge, and the technology to remain stable and present in the archaeological record. Levallois technique, then, offers the most likely example of a technology that requires teaching, and with it third level intentionality built on the linguistic skills that include such complex language abilities as the ability to describe mental states, and the linguistic structure to attribute them such as complementation. Levallois technique, then, while perhaps not requiring language to teach, requires a teacher that has language. In order to be a teacher, they must be linguistic to the point where they have developed their cognitive ability within a community of grammatical hominins. Figure 6.2 shows a broad (and necessarily reductive) schematic showing rough correlations between broad lithic industries and interpreted social learning techniques. It also offers the potential for progression where the social learning mechanisms selected for through one technology provide the context that facilitates the transition to the next. While at this resolution the model must necessarily be reductive, it also does not account for variation in cognitive ability and lithic behaviour represented by a varied and complex hominin phylogeny. But it provides a rough guide as to the broad signifying lithic behaviours and the cognitive and linguistic interpretations made at an industry level supported by the results of the experiment conducted for this thesis.

6.6 Summary

This chapter was the culmination of the results presented in Chapter 5 and the information presented in the earlier chapters. It reminded the reader of the purpose and motivation for the project in terms of providing an alternate proxy for Palaeolithic language ability than symbolism. It summarised the results and compared them to similar studies in the discipline to evaluate how in line the conclusions were. It also explored the interpretive potential of this thesis and highlighted its limitations. It then presented broad scale lithic industries in light of the cognitive and linguistic interpretations that this thesis lends, which is perhaps offers the clearest example of how these ideas can pertain to the archaeological record. What results is a novel and innovative approach to tracking the evolution of language in the Palaeolithic, and a powerful method for gathering supported interpretations of hominin cognitive and linguistic behaviour that takes us away from the conventional focus on symbolic material culture.

Chapter Seven.

Conclusion

7.1 An Assessment of Thesis Impact

This thesis has attained a number of methodological and theoretical accomplishments. It provides a method for the interpretation of historically elusive mental characteristics of ancient hominins. The challenge has been to take the perceived immateriality of the mind and connect it more saliently to the material world. Understanding material as fundamental to all, and the mind as both extended and distributed, allows for a paradigm that acknowledges how humans think through their tools, not just with them; this intimate connection between the material environment and cognition transforms the oft perceived gap to a smooth continuity between material and thought, between tools and their knowledge transmission.

The primary goal of this thesis has been to create a method that allows for the interpretation of linguistic ability from archaeological material. It looked to stone tools because of their pervasiveness, and their development throughout the timeframe in which language was evolving. Another proxy, symbolism, the most commonly cited link to cognitive ability, presents limitations in its recent timeframe, localised geography, poor connections to underlying mechanisms, and difficult detection in the record (especially before the Upper Palaeolithic). As language is a multifaceted ability, symbolism, being a binary attribute, offers only an indication or non-indication of the presence of linguistic ability, with no nuance. This presents serious issues for its use as a functional proxy that would illustrate the most likely cognitive evolutionary history of hominins. In this thesis, theory of mind was sought as a bridge between expressions of material culture that allude to behavioural complexity, and linguistic behaviours that necessitate those behavioural abilities. Theory of mind is interesting as a cognitive interpretation itself, but also insinuates so much more in terms of a behavioural life, including linguistic abilities for its development. This bridge connects when theory of mind can be identified in archaeological behavioural signatures, such as in complex social learning. Luckily, lithic assemblages provide a perfect opportunity to analyse what happens during the replication process in terms of material culture that has been transmitted through social learning. The variability of lithic tools is affected by a number of factors, including but by no means limited to: function, raw material, and skill. But as has been shown in this thesis, through the literature and also in the experiment, that copy error is introduced by the physical limitations on reproducing materials; copy error is shown to be introduced in varying amounts by the fidelity of the social learning method involved in the knowledge transmission. This thesis' experiment also supports that since low copy error will result in less variation, assemblages which show high levels of standardization must have been influenced less by features that introduce variation, and this includes the more minimal copy error introduced with high fidelity teaching methods. Therefore, highly standardised lithic assemblages could only have been produced in instances where there was high fidelity cultural transmission such as imitation or teaching behaviours. Being able to understand the different factors involved in assemblage variability and standardization can aid in constructing cognitive inferences based on the necessities of behaviour for that social transmission: that would be theory of mind, and the linguistic skills that scaffolded that development and allowed that cultural transmission to have taken place. In this case, the research question of the thesis has been answered.

7.2 Future Directions for Research

The use of porcelain in an experimental knapping setting allowed for the scientific recording of not only the target replications made by the participants, but the debitage that they created in its production. It allowed an in-depth morphological analysis of a kind not found in other knapping and social learning studies, and offered replicable methodologies for future studies' use. Use of the porcelain mimicked flint in the way the garden foam in Schillinger et al. (2015)'s study did not, allowed for a debitage analysis, and avoiding the pitfalls in raw material availability and variability faced by Ohnuma et al. (1997) and Putt et al. (2014).

The data in this study was analysed using a mixture of two complementary analytical approaches: a number of traditional measurements (weight loss, scar count and scar density, flake attributes) were used because it was hypothesised they would exhibit different relative ranges of morphological variation (as per this thesis' hypothesis). As well, a geometric morphometric methodology was employed to look at the shape variation of the outlines of three perspectives of the handaxes. The inclusion of the lateral and superior views of the handaxes gave a fuller, more global view of the shape differences between the groups, as opposed to the conventional analysis of just planar handaxe shape. This also allowed for a 3D perspective, without the time, equipment and expense that 3D scanning and analyses would require.

Throughout this thesis' experiment, features of the experimental design were highlighted that could be built upon for further study. In the experiment, the features of knapping that the participants needed to learn and replicate were all skills that could be acquired rather quickly and through a visual medium; this reduction in technical skill requirements was to isolate the variables learned to just a small few for analysis, but also out of necessity for thesis scope, time and financial constraints. As the handaxe processes could all be visually understood and replicated, it meant that the knapping features could all be acquired relatively easy by imitation and without explicit teaching, although the results showed that explicit teaching tended to increase the faithfulness of the replicas. As the success of Group 1 (Emulation) demonstrates, much of these processes could be reverse engineered from the end-state relatively faithfully.

All participants needed to learn how to strike a hammerstone against the porcelain blank to remove a series of flakes on both sides, and to shape one end to a tip. Participants came to the task with a prior knowledge that the hammerstone would be striking the blank to remove material, and so the unsolved connections were few. But what of techniques that appear in knapping behaviour archaeologically that are less intuitive, both in function and in mechanics? These techniques have been called 'opaque variables', as opposed to the more transparent variables that the participants needed to learn in the current study. Opaque variables could include preparing an edge for a removal with an abrasive surface in order to prevent diffusion of the energy through brittle material on the edge, which contributes to a thinner and more precise removal; or to make a series of small flakes on an edge in order to change the edge angle and remove a desired area of core material (important in core shaping). In prepared core techniques, where flakes are taken off in a planned manner in anticipation of a preferential flake, these behaviours could be considered opaque to the imitator; they require forward planning, and to an individual watching this activity without instruction or conferred insight, understanding the reasoning that the knapper has for making the decisions that they do could be quite non-evident. While bodily movements may be replicated, or over-imitated blindly, the purpose for abrading an edge or other skilled knapping strategy (unless it is made explicit by demonstration or explanation, both of which fall into the category of teaching) would likely not be learned and therefore not transmit and persist in the lithic record. If these opaque variables described here are those that might only be transmitted by high fidelity learning, there may be reason to extend this study to one which includes these techniques, and a more stark contrast between imitative and teaching learning environments' results may be seen.

Another area for future study could be in increasing the number of participants and number of handaxes each participant makes, extending the study into a more longitudinal format, so that participants are able to achieve a higher level of knapping expertise in their specialist social learning environment. This would allow the analysis of whether skill levels plateau in certain situations. The smaller number of handaxes made by each participant in the present study meant that an analysis of solely the final knapping attempts was not undertaken. Omitting the first few attempts (which were likely more variable due to the initial learning experiences) was not possible due to assemblage numbers. With more participants each having more attempts, the data could be broken down and analysed further, with high enough assemblage numbers to support robust statistics.

A further area of study could also to be to explore the effect of iterated 'generations' of knappers in each simulated social learning environment, such as that done in Morgan et al. (2015). The transmission could be further explored with the impact of multiple transmission points. This type of study requires an exponential amount of participants and materials, and would be something to explore when more resources were available and perhaps other aspects of the study and its implications were articulated further.

The research question of this thesis was directed at exploring variability in replication, and the experiment featuring handaxes was a matter of convenience and applicability. The goal was not to explore the cognitive requirements of replicating specifically handaxe technology. Indeed, the variables included in culturally transmitting the technological components of handaxe technology were dramatically reduced in the present study, and this was intentional. However, the design of the study could be modified to explore these features with a view to understanding the transmission of handaxe technology in particular. Acheulean technology spanned three continents over 1.5 million years and has been the focus of considerable debate concerning its persistence of form in the record. A modification of the current experiment to explore the transmission of specific features of Acheulean technology under different simulated social learning environments could be a way to investigate the necessary social learning methods required for its persistence in the record. By changing the target replication and features of the technology, this could equally be applied to other types of Palaeolithic technology (such as the transmission of creating elongated materials, or blades), and could even be applied beyond the necessities to transmit lithic technologies.

This thesis' study and its 2D shape analyses used different perspectives to approximate a 3D analysis. A full 3D analysis with appropriate scanning equipment and landmarks spread over the surfaces of the tools would provide a more accurate and higher resolution dataset to complement the analyses undertaken here.

7.3 A Final Note

Handaxes were once thought by some to be the product of thunderbolts (Roe 2003); by the 19th century, with the interest of antiquarians and an understanding of their deep antiquity, Brandon flint knappers could see the human agency in the handaxes because of their unique interaction with the material, and could reverse engineer the design. Forgeries started to surface to be sold to interested antiquarians, who now understood and valued their human agency and place in deep time. However, the forgers did not replicate the handaxes faithfully enough, and archaeologists began to recognise their modernity. A number of these forgeries exist in the Ashmolean museum, illustrating early archaeology's deep experimental roots. They used their metal tools (which they used to make gun flints), and rolled the handaxes in manure to give a patina, and so the forgeries are discernible by the angle and depth of the flake scars, and the type of staining. These Victorian flintknappers were emulating Palaeolithic handaxes, reverse engineering the process to create the forms that antiquarians were interested in. Today, there are communities of self-taught flint knappers, as well as those who have learned through those teachers. Emulative, imitative and teaching abilities are alive and well in the production of modern stone tools.

Speaking about stone tools in general, Deutscher (2005, p. 13) remarks that "there is still no compelling reason why these flaked stones could not have been produced without language". This thesis hopes to have provided a very compelling reason why they indeed required language – to have language, you need the cognitive scaffolding provided by theory of mind, and theory of mind is also necessary for the propagation and persistence of these flaked stones. The stones are cognitive fossils that once took part in a linguistic hominin's mental processes; with a method of identifying theory of mind and language ability in the archaeological record, a greater picture of hominin lifeways is opened up, and a robust understanding of Palaeolithic cultures can be accessed through their tools.

References

Abu-Akel, A. 2003. A neurobiological mapping of theory of mind. *Brain Research Reviews* 43, 29-40.

Aitchison, J. 2008. *The Articulate Mammal: an Introduction to Psycholinguistics.* London: Routledge.

Aplin, L. M., Sheldon, B. C., Morand-Ferron, J. 2013. Milk bottles revisited: social learning and individual variation in the blue tit, *Cyanistes caeruleus*. *Animal Behaviour* 85, 1225-1232.

Armstrong, D. F., and Wilcox, S. 2007. *The Gestural Origin of Language*. Oxford: Oxford University Press.

Asfaw, B., Beyene, Y., Suwa, G., Walter, R. C., White, T. D., WoldeGabriel, G., Yemane, T. 1992. The earliest Acheulean from Konso-Gardula. *Nature* 360, 732-735.

Astington, J. W. and Baird, J. A. 2005. *Why language matters for theory of mind.* New York: Oxford University Press.

Astington, J. W. and Jenkins, J. M. 1999. A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology* 35 (5), 1311-1320.

Astington, J. A., Pelletier, J., Homer, B. 2002. Theory of mind and epistemological development: the relation between children's second-order false-belief understanding and their ability to reason about evidence. *New Ideas in Psychology* 20, 131-144.

Avargues-Weber, A., and Chittka, L. 2014. Local enhancement or stimulus enhancement? Bumblebee social learning results in a specific pattern of flower preference. *Animal Behaviour* 97, 185-191.

Avis, J., and Harris, P. L. 1991. Belief-desire reasoning among Baka children: Evidence for a universal conception of mind. *Child Development* 62, 460-467.

Baldwin, D. A., and Moses, L. J. 1994. Early Understanding of Referential Intent and Attentional Focus: Evidence from Language and Emotion. In: Charles Lewis and Peter Mitchell (Eds.) *Children's Early Understanding of Mind: Origins and Development.* (pp. 133-156). Hove: Lawrence Erlbaum Associates.

Bar-Yosef, O. 2002. The Upper Palaeolithic Revolution. *Annual Review of Anthropology* 31, 363-393.

Barksy, D. 2009. An overview of some African and Eurasian Oldowan sites: evaluation of hominin cognition levels, technological advancement, and adaptive skills. In: Hovers, E. and Braun, D. R. (Eds.) *Interdisciplinary Approaches to the Oldowan*. Springer Netherlands.

Baron-Cohen, S. 1995. *Mindblindness: An essay on autism and theory of mind*. Boston: MIT Press/Bradford Books.

Baron-Cohen, S. and Cross, P. 1992. Reading the eyes: evidence for the role of perception in the development of a theory of mind. *Mind and Language* 7, 172-186.

Baron-Cohen, S., Jollife, T., Mortimore, C., Robertson, M. 1997. Another advanced test of theory of mind: evidence from very high functioning adults with autism or Asperger Syndrome. *The Journal of Child Psychology and Psychiatry* 38, 813-822.

Baron-Cohen, S., Leslie, A. M., & Frith, U. 1985. Does the autistic child have a "theory of mind"? *Cognition*, 21, 37–46.

Baron-Cohen, S., O'Riordan, M., Jones, R., Stone, V., & Plaisted, K. 1999. A new test of social sensitivity: Detection of faux pas in normal children and children with Asperger syndrome. *Journal of Autism and Developmental Disorders* 29, 407–418.

Baron-Cohen, S. and Wheelwright, S. 2004. The empathy quotient: an investigation of adults with Asperger Syndrome or High Functioning Autism, and normal sex differences. *Journal of Autism and Developmental Disorders* 34 (2), 163-175.

Bartsch, K. and Wellman, H. M. 1995. *Children talk about the mind*. New York: Oxford University Press.

Bednarik, R. G., 2008. Beads and Cognitive Evolution. Time and Mind: the Journal of Archaeology, *Consciousness and culture* 1 (3), 285-317.

Bennett, T., Szatmari, P., Bryson, S., Volden, J., Zwaigenbaum, L., Vaccarella, L. 2008. Differentiating autism and Asperger syndrome on the basis of language delay or impairment. *Journal of Autism and Developmental Disorders* 38, 616–625.

Bettinger, R. L. and Eerkens, J. 1999. Point typologies, cultural transmission, and the spread of bow-an-arrow technology in the prehistoric Great Basin. American Antiquity 64, 231-242.

Bickerton, D. 2009. *Adam's tongue: how humans made language, how language made humans.* New York: Hill and Wang.

Biro, D., Inoue-Nakamura, N., Tonooka, R., Yamakoshi, G., Sousa, C., Matsuzawa, T. 2003. Cultural innovation and transmission of tool use in wild chimpanzees: evidence from field experiments. *Animal Cognition* 6, 213-223.

Bora, E., Vahip, S., Gonoul, A.S., Akdeniz, F., Alkan, M., Ogut, M., & Eryavuz, A. 2005. Evidence for theory of mind deficits in euthymic patients with bipolar disorder. *Acta Psychiatrica Scandinavica*, 112, 110-116.

Botha, R. 2008. Prehistoric Shell Beads as a Window on Language Evolution. *Language and Communication* 28, 197-212.

Botha, R. 2010. On the Soundness of Inferring Modern Language from Symbolic Behaviour. *Cambridge Archaeological Journal* 20, 345-356.

Botha, R. and Everaert, M. 2013. Introduction: evidence and inference in the study of language evolution. In: Botha, R. and Everaert, M. (Eds.) *The Evolutionary Emergence of Language: Evidence and Inference*. Oxford: Oxford University Press.

Bowler, D. M., and Strom. E. 1998. Elicitation of first-order theory of mind in children with autism. *Autism* 2, 33-44.

Boyd, R. and Richerson, P. J. 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago.

Braun, D. R., and Hovers, E. 2009. Introduction: current issues in Oldowan research. In: Hovers, E. and Braun, D. R. (Eds.) *Interdisciplinary approaches to the Oldowan*. Springer Netherlands.

Bretherton, I., McNew, S., and Beeghly-Smith, M. 1981. Early person knowledge as expressed in gestural and verbal communication: when do infants acquire a "theory of mind"? In: M. E. Lamb and L. R. Sherrod (Eds.) *Infant Social Cognition* (pp. 333-373). Hillsdale, NJ: Erlbaum.

Brooks, R. and Meltzoff, A. N. 2008. Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: a longitudinal, growth curve modelling study. *Journal of Child Language* 35, 207-220.

Brühne, M. and Brühne-Cohrs, U. 2006. Theory of mind – evolution, ontogeny, brain mechanisms and psychopathy. *Neuroscience and Biobehavioural Reviews* 30, 437-455.

Bruinsma, Y., Koegel, R. L., and Koegel, L. K. 2004. Joint attention and children with autism: A review of the literature. *Mental Retardation and Developmental Disabilities*, 10, 169–175.

Burdukiewicz, J. M. 2014. The origin of symbolic behaviour of Middle Palaeolithic: recent controversies. *Quaternary International* 326-327, 398-405.

Burling , R., Armstrong, D. F., Blount, B. G., Callaghan, C. A., Foster, M. L., King, B. J., Parker, S. T., Osamu S., Stokoe, W. C., Wallace, R., Wallman, J., Whiten, A., Wilcox, S., and Wynn, T. 1993. Primate Calls, Human Language, and Nonverbal Communication [and Comments and Reply]. *Current Anthropology* 34, 25-53.

Burling, R. 2005. *The Talking Ape: How Language Evolved.* Oxford: Oxford University Press.

Buttelmann, D., Carpenter, M., Call, J., and Tomasello, M. 2007. *Developmental Science* 10 (4), F31-F38.

Buttelmann, D., Buttelmann, F., Carpenter, M., Call, J. and Tomasello, M. 2017. Great apes distinguish true from false beliefs in an interactive helping task. *PLoS ONE* 12,

Byrne, R. and Whiten, A., 1988. *Machiavellian Intelligence: social expertise and the evolution of intellect in monkeys, apes and humans.* Oxford, UK: Oxford University Press.

Call, J. and Tomasello, M., 1999. A nonverbal false belief task: the performance of children and great apes. *Child Development* 70, 381-395.

Call, J. and Tomasello, M., 2008. Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences* 12 (5), 187-192.

Caro, T. M. and Hauser, M. D. 1992. Is there teaching in nonhuman animals? *The Quarterly Review of Biology* 67, 151-174.

Carpenter, M. Nagell, K., Tomasello, M., Butterworth, G., and Moore, C., 1998. Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs for the Society for Research in Child Development* 63, 1-143.

Carrasco, L., Posada, S., Collel, M., 2009. New Evidence on Imitation in an enculturated chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology* 123 (4), 385-390.

Carruthers, P. and Smith, P. K. 1996. *Theories of theories of mind.* Cambridge: Cambridge University Press.

Cavalli-Sforza, L. L., and Feldman, M. W. 1981. *Cultural transmission and evolution: a quantitative approach.* Princeton: Princeton University Press.

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., and Drew, A. 2000. Testing joint attention, imitation and play as infancy precursors to language and theory of mind. *Cognitive Development* 15, 481-498.

Chase, P. G., and Dibble, H. L. 1987. Middle Palaeolithic symbolism: a review of current evidence and interpretations. *Journal of Anthropological Archaeology* 6, 263-296.

Chasiotis, A., Kiessling, F., Campos, D., and Hofer, J. 2006 Theory of mind and inhibitory control in three cultures: Conflict inhibition predicts false belief understanding in Germany, Costa Rica and Cameroon. *International Journal of Behavioral Development* 30, 249–260

Chazan, M. 2012. Handaxes, concepts and teaching. *Mind, Brain and Education* 6, 197-203.

Christiansen, M. H., and Kirby, S. 2003. Language evolution: consensus and controversies. *Trends in Cognitive Sciences* 7, 300-307.

Clark, G. A. 2009. Accidents of history: conceptual frameworks in Palaeoarchaeology. In: Camps, M. and Chauhan, P. (Eds.) *Sourcebook of Palaeolithic Transitions.* Springer: New York.

Clegg, J., Hollis, C., Mawhood, L., and Rutter, M. 2005. Developmental Language Disorders – a follow-up in later adult life. Cognitive, language and psychosocial outcomes. *The Journal of Child Psychology and Psychiatry* 46, 128-149.

Cole, J. Handaxe symmetry in the Lower and Middle Palaeolithic: implications for the Acheulean gaze. In: Coward, F., Hosfield, R., Pope, M. and Wenban-Smith, F. (Eds.) *Settlement, Society and Cognition in Human Evolution: Landscapes in Mind.* Cambridge: Cambridge University Press.

Corballis, M. C., 2012. The wandering mind: mental time travel, theory of mind, and language. *Análise Social* 205, 870-893.

Corbey, R., Jagich, A., Vaesen, K. and Collard, M. 2016. The acheulean handaxe: more like a bird's song than a Beatles' tune? *Evolutionary Anthropology* 25, 6-19.

Costa, A. G. 2010. A geometric morphometric assessment of plan shape in bone and stone Acheulean bifaces from the Middle Pleistocene site of Castel di Guido, Latium, Italy. In: Lycett, S. J., and Chauhan, P. R. (Eds.) *New Perspectives in Old Stones*. New York: Springer.

Courtin, C. 2000. The impact of sign language on the cognitive development of deaf children: the case of theories of mind. *Journal of Deaf Studies and Deaf Education* 5 (3) 266-276.

Craparo, R. M. 2007. Significance level. In: Salkind, N. J. (Ed.) Encyclopedia of Measurement and Statistics. Thousand Oaks, CA: SAGE Publications. Pp. 889-891.

Crystal, D. 1995. *The Cambridge encyclopedia of the English language*. Cambridge: Cambridge University Press.

Cutting, A. L. and Dunn, J. 1999. Theory of mind, emotion understanding, language, and family background: individual differences and interrelations. *Child Development* 70 (4), 853-865.

d'Errico, F., Henshilwood, C., Lawson, G., Vanhaeren, M., Tillier, A.-M., Soressi, M., Bresson, F., Maureille, B., Nowell, A., Lakarra, J., Backwell, L., and Julien, M. 2003. Archaeological evidence for the emergence of language, symbolism, and music—an alternative multidisciplinary perspective. *Journal of World Prehistory* 17, 1-70.

de Saussure, F., Bally, C., Sechehaye, A., and Riedinger, A. Cours de linguistique générale. Paris, Lausanne.

de Villiers, J. 2005. Can language acquisition give children a point of view? In: Astington, J. W., and Baird, J. A. (Eds.) *Why Language Matters for Theory of Mind.* (pp. 186-219). New York: Oxford University Press.

de Villiers, J. 2007. *The interface of language and theory of mind*. *Lingua* 117 (11), 1858-1878.

de Villiers, J., Pyers J. E. 2002. Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development* 17, 1037–1060.

Dahlgren, S., Dahlgren Sandberg, A., Larsson, M. 2010. Theory of mind in children with severe speech and physical impairments. *Research in Developmental Disabilities* 31, 617-624.

Dawkins, R. and Krebs, J. R. 1978. Animal signals: information or manipulation? In J. R. Krebs and N. B. Davies (Eds.) *Behavioural Ecology* (pp. 282-309) Oxford: Blackwell Scientific Publications.

Davidson, I. 1991. The archaeology of language origins – a review. *Antiquity* 65, 39-48.

Delagnes, A., and Roche, H. 2005. Late Pliocene hominid knapping skills: the case of Lokalalei 2C, West Turkana, Kenya. *Journal of Human Evolution* 48, 435-372.

Dennett, D. C. 1983. Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. *Behavioral and Brain Sciences* 6, 343-355.

Deutscher, G. 2005. *The unfolding of language.* London: Arrow Books.

Doherty, M. 2009. *Theory of mind: how children understand others' thoughts and feelings.* New York: Psychology Press.

Dunbar, R. I. M. 1998. Theory of mind and the evolution of language. In Hurford, J. R. and Studdert-Kennedy, M. and Knight C. (Eds.) *Approaches to the Evolution of Language: Social and Cognitive Bases.* (pp. 92-110) Cambridge: Cambridge University Press.

Dunbar, R. I. M., 2003. The Social Brain: Mind, Language and Society in Evolutionary Perspective. *Annual Review of Anthropology* 32, 163-181.

Dunbar, R. I. M. 2004. *The human story: a new history of mankind's evolution.* London: Faber and Faber.

Ebert, S. 2015. Longitudinal relations between theory of mind and metacognition and the impact of language. *Journal of Cognition and Development* 16, 559-586.

Eerkens, J.W., and Lipo, C.P. 2005. Cultural transmission, copying errors, and the generation of variation in material culture and the archaeological record. *Journal of Anthropological Archaeology*, 24, 316-334.

Eerkens, J. W. and Lipo, C. P. 2007. Cultural transmission theory and the archaeological record: providing context to understanding variation and temporal changes in material culture. *Journal of Archaeological Research* 15, 239-274.

Eisenmajer, R and Prior, M. 1991. Cognitive linguistic correlates of 'theory of mind' ability in autistic children. *British Journal of Developmental Psychology* 9, 351-364.

Ellen, R. Lycett, S. J., Johns, S. E. 2013. *Understanding Cultural Transmission in Anthropology: a Critical Synthesis.* New York: Berghahn Books.

Evans, J. 1861. On the occurrence of flint implements in undisturbed beds of gravel, sand and clay. *Archaeologica* 38, 280-307.

Farrar, M. J., and Maag, L., 2002. Early language development and the emergence of a theory of mind. *First Language* 22, 197–213.

Farrar, M. J., Johnson, B., Tompkins, V., Easters, M., Zilisi-Medus, A., and Benigno, J. P. 2009. Language and theory of mind in preschool children with specific language impairment. *Journal of Communication Disorders* 42, 428-441.

Fay, M. P., and Proschan, M. A., 2010. Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Statistics Surveys* 4, 1-39.

Fisher, J. and Hinde, R.A. 1949. The opening of milk bottles by birds. *British Birds* 42, 347-357.

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

Flavell, J. H., Everett, B. A., Croft, K., and Flavell, E. R. 1981. Young children's knowledge about visual perception: further evidence for the level 1-level 2 distinction. *Developmental Psychology* 17, 99-103.

Frith, U. and Frith, C. D. 2003. Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London B* 358, 459-473.

Fritz, J., Bisenberger, A., Kotrschal, K. 2000. Stimulus enhancement in greylag geese: socially mediated learning of an operant task. *Animal Behaviour* 59, 1119-1125.

Frye, D. and Ziv, M. 2004. Children's understanding of teaching: the role of belief and knowledge. *Cognitive Development* 19, 457-477.

Garfield, J. L., Peterson, C. C., and Tricia P. 2001. Social cognition, language acquisition and the development of theory of mind. *Mind and Language* 16, 494-541.

Gelman, R. and Spelke, E. 1981. The development of thoughts about animate and inanimate objects: implications for research on social cognition. In: Flavell, J. and Ross, L. (Eds.) *Social Cognitive Development: Frontiers and Possible Futures*. New York: Cambridge University Press.

Gerrans, P. 2002. The theory of mind module in evolutionary psychology. *Biology and Philosophy* 17, 305-321.

Gómez, J. C. 2007. Pointing behaviors in apes and human infants: a balanced interpretation. *Child Development* 78, 729-734.

Gordon, R. M. 1986. Folk psychology as simulation. *Mind and Language* 1, 158-171.

Hale C.M., Tager-Flusberg, H. 2003. The influence of language on theory of mind: A training study. *Developmental Science* 6, 346–359.

Hammer, Ø. And Harper, D. A. T. 2006. *Palaeontological Data Analysis*. Blackwell Publishing.

Hare, B., Call, J., Agnetta, B., and Tomasello, M. 2000. Chimpanzees know what conspecifics do and do not see. *Animal Behaviour* 59, 771-785.

Harmand, S., Lewis, J. E., Feibel, C. S., Lepre, C. J., Prat, S., Lenoble, A., Boës, X., Quinn, R. L., Brenet, M., Arroyo, A., Taylor, N., Clément, S., Daver, G., Brugal, J.-P., Leakey, L., Mortlock, R. A., Wright, J. D., Lokorodi, S., Kirwa, C., Kent, D. V., and Roche, H.2015. 3.3-million-year-old stone tools from Lomekwi 3, West Turkana, Kenya. *Nature* 521, 310-315.

Hamilton, A., Brindley, R., and Frith, U. 2009. Visual perspective taking impairment in children with autistic spectrum disorder. *Cognition* 113, 37-44.

Hammer, Ø., Harper, D.A.T. and Ryan, P.D. (2001). PAST: Paleontological statistics package for education and data analysis. *Palaeontologia Electronica*, 4, 1-9.

Happé, F. G. E. 1994. An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders* 24, 129-154.

Happé, F. G. E. 1995. The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development* 66, 843-855.

Hao, J., Su, Y., and Chan, R. C. K. 2010. Do deaf adults with limited language have advanced theory of mind? *Research in Developmental Disabilities* 31, 1491-1501.

Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., Chomsky, N., and Lewontin, R. C. 2014. The mystery of language evolution. *Frontiers in Psychology* 5 (401), 1-12.

Hayden, B. 2015. Insights into early lithic technologies from ethnography. *Philosophical Transactions of the Royal Society B* 370, 20140356.

Henshilwood, C. S., and Dubreuil, B. 2009. Reading the artifacts: gleaning language skills from the Middle Stone Age in southern Africa. In: Botha, R. and Knight, C. (Eds.) *The Cradle of Language.* Oxford: Oxford University Press, 41-61.

Henshilwood, C. S., and Marean, C. W. 2003. The origin of modern human behaviour: critique of the models and their test implications. *Current Anthropology* 44, 627-651.

Heyes, C. M. 1994. Social learning in animals: categories and mechanisms. *Biological Reviews* 69, 207-231.

Heyes, C. M. and Galef, B. G. (Eds.) 1996. *Social learning in animals: The roots of culture.* New York: Academic Press.

Hopkinson, T., Nowell, A., and White, M. 2013. Life histories, Metapopulation Ecology, and Innovation in the Acheulean. *Palaeoanthropology* 2013, 61-76.

Hopper, L. M., Spiteri, A., Lambeth, S. P., Schapiro, S. J., Horner, V., and Whiten, A. 2007. Experimental studies of traditions and underlying transmission processes in chimpanzees. *Animal Behaviour* 73, 1021-1032.

Horne, P. J., Erjavec, M., Lovett, V. E. 2008. The effects of modelling, local stimulus enhancement, and affordance demonstration on the production of object-directed actions in 6-month-old infants. *British Journal of Developmental Psychology* 27, 269-281.

Hovers, E. 2012. The makings of Oldowan lithic technology. In: Elias, S. (Ed.) *Origins of Human Innovation and Creativity*. Oxford: Elsevier.

Huang, C.-T., and Charman, T. 2005. Gradations of emulation learning in infants' imitation of actions on objects. *Journal of Experimental Child Psychology* 92. 276-302.

Hughes, C. 1998. Executive function in pre-schoolers: links with theory of mind and verbal ability. *British Journal of Developmental Psychology* 16, 233-253.

Hughes, C., and Dunn, J. 1997. "Pretend you didn't know": pre-schoolers talk about mental states in pretend play. *Cognitive Development* 12, 477-499.

Iovita, R., Tuvi-Arad I., Moncel, M.-H., Desiprée J., Voinchet, P. and Bahain J.-J. 2017. High handaxe symmetry at the beginning of the European Acheulean. The data from La Noira (France) in context. *PLoS ONE*: 12, e0177063.

Jenkins, J. M., and Astington, J. W. 1996. Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology* 32, 70-78.

Kaminski, J., Call, J., Tomasello, M. 2008. Chimpanzees know what others know, but not what they believe. *Cognition* 109, 224-234.

Kawai, M. 1965. Newly-acquired pre-cultural behaviour of the natural troop of Japanese monkeys on Koshima Islet. *Primates* 6, 1-30.

Kempe, M., Lycett, S. and Mesoudi, A. 2012. An experimental test of the accumulated copy error model of cultural mutation for Acheulean handaxe size. *PLoS ONE* 7, e48333.

Khreisheh, N. N., Davies, D., Bradley, B. A. 2013. Extending Experimental Control: the use of porcelain in flaked stone experimentation. *Advances in Archaeological Practice: A Journal of the Society for American Archaeology* 1(1), 37–46.

Kinsella, A. 2009. *Language Evolution and Syntactic Theory*. Cambridge: Cambridge University Press.

Kjelgaard, M. M., and Tager-Flusberg, H. 2001. An investigation of language impairments in autism: Implications for genetic subgroups. *Language and Cognitive Processes* 16, 287–308.

Klemenic, S. 2012. *An experimental investigation into the Clovis snapped blade technology*. Unpublished Bachelor's thesis, University of Exeter.

Korkmaz, B. 2011. Theory of Mind and Neurodevelopmental Disorders of Childhood. *Pediatric Research* 69, 101R-108R.

Kováč, L. 2015. *Closing Human Evolution: Life in the Ultimate Age*. Springer International Publishing.

Krachun, C., Carpenter, M., Call, J., and Tomasello, M., 2009. A competitive nonverbal false belief task for children and apes. *Developmental Science* 12, 521-535.

Kruger, A. C., and Tomasello, M. 1996. Cultural learning and learning culture. In: D. R. Olson and N. Torrance (Eds.), *Handbook of education and human development: New models of learning, teaching, and schooling* (pp. 369-387). Cambridge: Blackwell.

Krupenye, C., Kano, F., Hirata, S., Call, J., and Tomasello, M. 2016, Great apes anticipate that other individuals will act according to false beliefs. *Science* 354, 110-113.

Langley, M. 2015. Symbolic material culture in human evolution: Use in prehistory, appearance in the archaeological record and taphonomy. In: B. Putová and V. Soukup (Eds.), *The Genesis of Creativity and the Origin of the Human Mind*. (pp. 57-75). Karolinum Press.

Leakey, M. D. 1971. *Olduvai Gorge III: excavations in beds I and II, 1960-1963.* Cambridge: Cambridge University Press.

Lefebvre, L. 1995. The opening of milk bottles by birds: evidence for accelerating learning rates, but against the wave of advance model of cultural transmission. *Behavioral Processes* 34, 43-53.

Lepre, C. J., Roche, H., Kent, D. V., Harmand, S., Quinn, R. L., Brugal, J.-P.,Texier, P.-J., Lenoble, A. and Feibel, C. An earlier origin for the Acheulean. *Nature* 477, 82-85.

Leslie, A. M. 1987. Pretense and representation: the origins of "theory of mind". *Psychological Review* 94, 412-426.

Leslie, A. M., Friedman, O., and German, T. P. Core mechanisms in 'theory of mind'. Trends in Cognitive Sciences 8, 528-533.

Leudar, I. and Costall, A. 2009. *Against theory of mind.* London: Palgrave Macmillan.

Liddle, B. and Nettle, D. 2006. Higher-order theory of mind and social competence in school-age children. *Journal of Cultural and Evolutionary Psychology* 4, 231-246.

Lihoreau, M. and Rivault, C. 2011. Local enhancement promotes cockroach feeding aggregations. *PLoS ONE* 6, e22048.

Lindly, J. M., and Clark, G. A. 1990. Symbolism and modern human origins. *Current Anthropology* 31, 233-61.

Liu, D., Meltzoff, A. N., Wellman, H. M. 2009. Neural correlates of belief- and desire-reasoning. *Child Development* 80, 1163-1171.

Lockl, K. and Schneider, W. 2007. Knowledge about the mind: Links between theory of mind and later metamemory. *Child Development* 78, 148-167.

Lohman, H. and Tomasello, M. 2003. The role of language in the development of false belief understanding : A training study. *Child Development* 74, 1130-1144.

Lycett, S. J. 2010. The importance of history in definitions of *culture*: implications from phylogenetic approaches to the study of social learning in chimpanzees. *Learning and Behavior* 38, 252-264.

Lycett, S. J. 2015. Cultural evolutionary approaches to artefact variation over time and space: basis, progress, and prospects. *Journal of Archaeological Science* 56, 21-31.

Lycett, S. J. and Gowlett, J. A. J. 2008. On questions surrounding the Acheulean 'tradition'. *World Archaeology* 40, 295-315.

Lycett, S. J., von Cramon Taubadel, N., and Foley, R. A. 2006. A crossbeam co-ordinate caliper for the morphometric analysis of lithic nuclei: a description, test, and empirical examples of application. *Journal of Archaeological Science* 33, 847-861.

Lycett, S. J., Schillinger, K., Kempe, M., Mesoudi, A. 2015. Learning in the Acheulean: Experimental insights using handaxe form as a 'model organism'. In: (Mesoudi, A. and Aoki, K., Eds.) *Learning Strategies and Cultural Evolution During the Palaeolithic.* Springer Japan. pp.155-166.

Lycett, S. J., Schillinger, K., Eren, M., von Cramon-Taubadel, N., and Mesoudi, A. 2016a. Factors affecting Acheulean handaxe variation: experimental insights, microevolutionary processes, and macroevolutionary outcomes. *Quaternary International* 411, 386-401.

Lycett, S. J., von Cramon-Taubadel, N., and Eren, M. I. 2016b. Levallois: potential implications for learning and cultural transmission capacities. *Lithic Technology* 41, 19-38.

Malafouris, L. 2013. *How things shape the mind: a theory of material engagement*. London: MIT Press.

Malle, B. F., 2002. The relation between language and theory of mind in development and evolution. In T. Givón and B. F. Malle (Eds.), *The evolution of language out of pre-language* (pp. 265-284). Amsterdam: Benjamins.

Marshall, G. D., Gamble, C. G., Roe, D., and Dupplaw, D. 2002. *Acheulean Biface Database.* URL: http://ads.ahds.ac.uk/catalogue/specColl/bifaces/bf_query.cfm. York: Archaeology Data Service.

Marshall-Pescini, S. and Whiten, A. 2008. Social learning of nut-cracking behaviour in East African sanctuary-living chimpanzees (Pan troglodytes schweinfurthii). *Journal of Comparative Psychology* 122, 186-194.

Mesoudi, A. and O'Brien, M. J. 2008. The cultural transmission of Great Basin projectile-point technology I: an experimental simulation. *American Antiquity* 73, 3-28.

Marshack, A. 1976. Some implications of the Palaeolithic symbolic evidence for the origin of language. *Annals of the New York Academy of Sciences* 280, 289-311.

Marticorena, D. C. W., Ruiz, A. M., Mukerji, C., Goddu, A., Santos, L. R. 2011. Monkeys represent others' knowledge but not their beliefs. *Developmental Science* 14, 1406-1416.

McBrearty, S. and Brooks, A. S. 2000. The revolution that wasn't: a new interpretation of the origin of modern human behaviour. *Journal of Human Evolution* 39, 453-563.

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

McNabb, J. and Cole, J. 2015. The mirror cracked: symmetry and refinement in the Acheulean handaxe. *Journal of Archaeological Science: Reports* 3, 100-111.

McPherron, S. P., and Dibble, H. L. 1999. Stone tool analysis using digitized images: examples from the Lower and Middle Palaeolithic. *Lithic Technology* 24, 38-52.

McPherron, S. P. 2000. Handaxes as a measure of the mental capacities of early hominids. *Journal of Archaeological Science* 27, 655-663.

Mellars, P. A. 1996. *The Neanderthal Legacy: an archaeological perspective from Western Europe*. Princeton: Princeton University Press.

Meltzoff, A. N. 1988. Infant imitation after a 1-week delay: long-term memory for novel acts and multiple stimuli. *Developmental Psychology* 24, 470-476.

Mesoudi, A., Whiten, A., and Laland, K. 2006. Towards a unified science of cultural evolution. *Behavioral and Brain Sciences* 29, 329-347.

Mesoudi, A. and Aoki, K. 2015. *Learning Strategies and Cultural Evolution during the Palaeolithic*. Springer Japan.

Mesoudi, A. and O'Brien, M. J. The cultural transmission of Great Basin projectile-point technology I: an experimental simulation. *American Antiquity* 73, 3-28.

Miller, C. A. 2006. Developmental relationships between Language and Theory of Mind. *American Journal of Speech Language Pathology* 15, 142-154.

Miller, S. A. 2009. Children's understanding of second-order mental states. *Psychological Bulletin* 135, 749-773.

Milligan, K., Astington, J. W., Dack, L. A. 2007. Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child Development* 78 (2), 622-646.

Monnier, G. 2006, Testing retouched flake tool standardization during the Middle Palaeolithic: patterns and implications. In: Hovers, E. and Kuhn, S. (Eds.) *Transitions Before the Transition*. Springer.

Morgan, T. J. H., Uomini, N., Rendall, L. E., Chouinard-Thuly, L., Street, S. E., Lewis, H. M., Cross, C, P., Evans, C., Kearney, R., de la Torre, I., Whiten, A., and Laland, K. N. 2015. Experimental evidence for the co-evolution of hominin tool-making, teaching and language. *Nature Communications* 6 (6029), 1-8.

Musgrave, S., Morgan, D., Lonsdorf, L., Mundry, R., and Sanz C. 2016. Tool transfers are a form of teaching in chimpanzees. *Nature: Scientific Reports* 6, 34783.

Myowa-Yamakoshi, M., & Matsuzawa, T. 2000. Imitation of intentional manipulatory actions in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology* 114, 381-391.

Ohnuma, K., Aoki, K., Akazawa, T. 1997. Transmission of tool-making through verbal and non-verbal communication: preliminary experiments in Levallois flake production. *Anthropological Sciences* 105 (3), 159-168.

O'Leary, S. 2012. Can we identify adzes in the archaeological record? An experimental study based on material from the Gault site, Texas. Unpublished Bachelor's Thesis, University of Exeter.

Olson, D. R., & Bruner, J. S. 1996. Folk psychology and folk pedagogy. In D. R. Olson & N. Torrance (Eds.), *The Handbook of Education and Human Development* (pp. 9–27). Oxford: Blackwell.

Oudeyer, P.-Y. 2005. The self-organization of speech sounds. *Journal of Theoretical Biology* 233, 435-449.

Page, S. 2014. Cultural transmission of lithic artefact traditions: an experimental approach. Unpublished PhD thesis: University College London.

Panger, M. A., Brooks, A. S., Richmond, B. G., Wood, B. 2002. Older than the Oldowan? Rethinking the emergence of hominin tool use. *Evolutionary Anthropology* 11, 235-245.

Paynter, J. and Peterson, C. 2010. Language and ToM development in autism versus Asperger syndrome: contrasting influences of syntactic versus lexical/semantic maturity. *Research in Autism Spectrum Disorders* 4, 377-385.

Penn, D. C., and Povinelli, D. J. 2007. On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Philosophical Transactions of the Royal Society B* 362, 731-744.

Perner, J. and Wimmer, H. 1985. "John thinks that Mary thinks that…": attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology* 39, 437-471.

Pfungst, O. 1911. *Clever Hans: The horse of Mr. Von Osten.* Oxford: Henry Holt and Company.

Premack, D. and Woodruff, G. 1979. Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences* 1978 (4), 515-526.

Putt, S. S., Woods A. D., and Franciscus R. G. 2014. The role of verbal interaction during experimental bifacial stone tool manufacture. *Lithic Technology* 39 (2), 96-112.

Rakhlin, N., Kornilov, S. A., Reich, J., Babyonyshev, M., Koposov, R. A., and Grigorenko, E. L. 2011. The relationship between syntactic development and theory of mind: evidence from a small-population study of a developmental language disorder. *Journal of Neurolinguistics* 24, 476-496.

Range, F., Viranyi, Z., and Huber, L. 2007. Selective imitation in domestic dogs. *Current Biology* 17, 868-872.

Repacholi, B. M. and Gopnik, A., 1997. Early reasoning about desires: evidence from 14- and 18-month-olds. *Developmental Psychology* 33, 12-21.

Reaux, J. E., Theall, L. A., Povinelli, D. J. 1999. A longitudinal investigation of chimpanzees' understanding of visual perception. *Child Development* 70, 275-290.

Remmel, E. and Peters, K. 2009. Theory of mind and language in cochlear implants. *Journal of Deaf Studies and Deaf Education* 14, 218-236.

Richerson, P. J., and Boyd, R. 2005. *Not by genes alone: how culture transformed human evolution.* Chicago: Chicago University Press.

Roberts, M. B. and Parfitt, S. A. 1999. *Boxgrove: a Middle Pleistocene Hominid Site at Eartham Quarry, Boxgrove, West Sussex.* London: English Heritage.

Roe, D. A. 2003. No stone unturned: a look at the study of stone tools in early prehistory. *Lithics* 24, 37-56.

Rohlf, F.J. (2016a). tpsUtil Program. Version 1.69. SUNY, Stony Brook.

Rohlf, F.J. (2016b). tpsDig2. Version 2.27. SUNY, Stony Brook.

Rohlf F. J. and Slice, D. 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Biology* 39, 40-59.b

Ruffman, T., Slade, L., Rowlandson, K., Rumsey, C. Garnham, A. 2003. How language relates to belief, desire and emotion understanding. *Cognitive Development* 18, 139-158.Schick, B., de Villiers, P., de Villiers, J., and Hoffmeister, R. 2007. Language and theory of mind: a study of deaf children. *Child Development* 78 (2), 376-396.

Schick, K.D., & Toth, N. 1993. *Making silent stones speak: human evolution and the dawn of technology*. London: Clays Ltd.

Schillinger, K., Mesoudi, A., & Lycett, S. J. 2014. Copying error and the cultural evolution of “additive” vs. “reductive” material traditions: an experimental assessment. *American Antiquity* 79, 128-143.

Schillinger, K., Mesoudi, A., Lycett, S. J. 2015. The impact of imitative versus emulative learning mechanisms on artifactual variation: implications for the evolution of material culture. *Evolution and Human Behavior* 36, 446-455.

Schlanger, N. 1996. Understanding Levallois: lithic technology and cognitive archaeology. *Cambridge Archaeological Journal* 6, 231-254.

Scholl, B. J. and Leslie, A. M. 1999. Modularity, development and ‘theory of mind’. *Mind and Language* 14, 131-153.

Segal, G. 1998. Representing representations. In: P. Carruthers and J. Boucher (Eds.), *Language and Thought: Interdisciplinary Themes* (pp. 146-161). Cambridge, Cambridge University Press.

Semaw, S. 2000. The world’s oldest stone artefacts from Gona, Ethiopia: their implications for understanding stone technology and patterns of human evolution between 2.6-1.5 million years ago. *Journal of Archaeological Science* 27, 1197-1214.

Shapiro, S. S. and Wilk, M. B. 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52, 591-611.

Shatz, M. 1994. *A Toddler's Life: Becoming a Person.* New York: Oxford University Press.

Sherry, D. F., and Galef, B. G. 1984. Cultural transmission without imitation: Milk bottle opening by birds. *Animal Behaviour* 32, 937-938.

Shipton, C. B. K., Petraglia, M. & Paddayya, K. 2009. Stone tool experiments and reduction methods at the Acheulean site of Isampur Quarry, India. *Antiquity* 83, 769–785.

Slade, L. and Ruffman, T. 2005. How language does (and does not) relate to theory of mind: a longitudinal study of syntax, semantics, working memory and false belief. *British Journal of Developmental Psychology* 23, 117-141.

Small, C. G. 1996. *The statistical theory of shape*. New York: Springer-Verlag.

Smith, P. K. 1996. Language and the evolution of mind-reading. In: Carruthers, P. and Smith, P. K. *Theories of theories of mind.* Cambridge: Cambridge University Press.

Solé, R. 2005. Language: Syntax for free? *Nature* 434, 289.

Sprung, M. 2010. Clinically relevant measures for children's theory of mind and knowledge about thinking: non-standard and advanced measures. *Child and Adolescent Mental Health* 15 (4), 204-216.Stade, C. M., and Gamble, C. In three minds: extending cognitive archaeology with the social brain. In: Squeezing minds from stones: cognitive archaeology and the evolution of the human mind. New York: Oxford University Press. *Accepted.*

Stark, M. T., Bowser, B. J., Horne, L. 2008. *Cultural Transmission and Material Culture: breaking down boundaries*. Tucson: University of Arizona Press.

Stich, S. and Nichols, S. 1992, Folk psychology: simulation or tacit theory? *Mind and Language* 7, 35-71.

Stout, D. 2002. Skill and cognition in stone tool production: an ethnographic case study from Irian Jaya. *Current Anthropology* 43, 693-722.

Stout, D., Apel, J., Commander, J., Roberts, M., 2014. Late Acheulean technology and cognition at Boxgrove, UK. *Journal of Archaeological Science* 41, 576-590.

Stout, D., Semaw, S., Rogers, M. J., Cauche, D. 2010. Technological variation in the earliest Oldowan from Gona, Afar, Ethiopia. *Journal of Human Evolution* 58, 474-491.

Strauss, S., Ziv, M., Stein, A. 2002. Teaching as a natural cognition and its relations to preschoolers' developing theory of mind. *Cognitive Development* 17, 1473-1787.

Sutton, J., Smith, P. K., Swettenham, J. 1999. Bullying and 'theory of mind': a critique of the 'social skills deficit' view of anti-social behaviour. *Social Development* 8, 117-127.

Tager-Flusberg, H., and Joseph, R. M. 2005. How Language Facilitates the Acquisition of False-Belief Understanding in Children with Autism. In: Astington, J. W., and Baird, J. A. (Eds.) *Why Language Matters for Theory of Mind* (pp 298-318). New York: Oxford University Press.

Tehrani, J. J., and Reide, F. 2008. Towards an archaeology of pedagogy: learning, teaching and the generation of material culture traditions. *World Archaeology* 40 (3), 316-331.

Tennie, C., Call, J., and Tomasello, M. 2009. Ratcheting up the ratchet: on the evolution of cumulative culture. *Philosophical Transactions of the Royal Society* B 364, 2405-2415.

Tennie, C., Call, J., and Tomasello, M. 2010. Evidence for emulation in chimpanzees in social settings using the floating peanut task. *PLoS ONE* 5: e10544.

Terkel, J. 1996. Cultural transmission of feeding behavior in the black rat (*Rattus rattus*). In: Heyes, C. M. and Galef, B. G (Eds.) *Social Learning in Animals: the Roots of Culture.* San Diego: Academic Press. pp. 17–47.

Träuble, B., Marinović, V., and Pauen, S. 2010. Early theory of mind competencies: do infants understand others' beliefs? *Infancy* 15, 434-444.

Tomasello, M. 1995. Joint attention as social cognition. In: C. Moore and P. Dunham (Eds.) *Joint Attention: its Origins and Role in Development.* New York: Psychology Press.

Tomasello, M. 1996. Do apes ape? In Heyes, C. M. and Galef, B. G. (Eds.), *Social learning in animals: The roots of culture* (pp. 319–346) New York: Academic Press.

Tomasello. M. and Call, J. 1997. *Primate Cognition*. New York: Oxford University Press.

Tomasello, M., Davis-Dasilva, M., and Camak, L. 1987. Observational Learning of tool-use by young chimpanzees. *Human Evolution* 2, 175-183.

Toth, N. 1985. The oldowan reassessed: a close look at early stone artifacts. Journal of *Archaeological Science* 12, 101-120.

Toth, K., Munson, J., Meltzoff, A. N. 2006. Early predictors of communication development in young children with Autism Spectrum Disorder: joint attention, imitation and toy play. *Journal of Autism and Developmental Disorders* 36, 993-1005.

Träuble, B., Marinović, V., and Pauen, S. 2010. Early theory of mind competencies: do infants understand others' beliefs? *Infancy* 15, 434-444.

Voekl, B. and Huber, L. 2000. True imitation in marmosets. *Animal Behaviour* 60, 195-202.

Vanhaeren, M., d'Errico, F., van Niekerk, K. L., Henshilwood, C. S., Erasmus, R. M. 2013. Thinking Strings: additional evidence for personal adornment use in the Middle Stone Age at Blombos Cave, South Africa. *Journal of Human Evolution* 64, 500-517.

Vogeley, K., Bussfeld, P., Newen, A., Herrman, S., Happé, F., Falkai, P., Maier, W., Shah, N. J., Fink, G. R., and Zilles, K. 2001. Mind reading: neural mechanisms of theory of mind and self-perspective. *NeuroImage* 14, 170-181.

Wasielewski, H. 2014. Imitation Is Necessary for Cumulative Cultural Evolution in an Unfamiliar, Opaque Task. *Human Nature* 25, 161-179.

Weinrich, M. T., Schilling, M. R. & Belt, C. R. 1992. Evidence for acquisition of a novel feeding behaviour: lobtail feeding in humpback whales, *Megaptera novaeangliae*. *Animal Behaviour* 44, 1059-1072.

Wellman, H. M., Cross, D. and Watson, J. 2001. Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development* 72, 655-684.

Wellman, H. M., Fung, K. Y. F., Liu, D., Fang, F. 2005. Preschoolers' understanding of knowing-that and knowing-how in the United States and Hong Kong. *Developmental Psychology* 41, 562-573.

Whiten, A. and Byrne, R. W. 1988. Tactical deception in primates. *Behavioral and Brain Sciences* 11, 233-273.

Whiten, A., Custance, D. M., Gomez, J.-C., Teixidor, P. and Bard, K. A. 1996. Imitative learning of artificial fruit processing in children (*Homo sapiens*) and chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology* 110, 3-14.

Whiten, A. Horner, V., Litchfield, C. A., Marshall-Pescini, S. 2004. How do apes ape? *Learning & Behavior* 32, 36-52.

Whiten, A., McGuigan, N., Marshall-Pescini, S., and Hopper, L. 2009. Emulation, imitation, and over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B* 364, 2417-2428.

Whittaker, J. C., 1994. *Flintknapping: making and understanding stone tools*. Austin: University of Texas Press.

Wimmer, H. and Perner, J., 1983. Beliefs about beliefs: representing and constraining function of wrong belief in children's understanding of deception. *Cognition* 13 (1), 103-128.

Worden, R. 1998. The evolution of language from social intelligence. In: J. R. Hurford, M. Studdert-Kennedy, C. Knight. *Approaches to the Evolution of Language.* Cambridge: Cambridge University Press. (pp. 148-166)

Wynn, T. 1995. Handaxe enigmas. *World Archaeology* 27, 10-24.

Wynn, T. 2009. Whither evolutionary cognitive archaeology? Afterword. In: de Beaune, S. A., Coolidge, F. L., Wynn, T. (Eds.) *Cognitive Archaeology and Human Evolution*. (pp. 145-149) New York: Cambridge University Press.

Wynn, T. and Coolidge, F. L. 2009. Implications of a strict standard for recognizing modern cognition in prehistory. In: de Beaune, S. A., Coolidge, F. L., Wynn, T. (Eds.) *Cognitive Archaeology and Human Evolution*. (pp. 117-128) New York: Cambridge University Press.

Wynn, T. and Coolidge, F. 2012. *How to think like a Neanderthal*. Oxford: Oxford University Press.

Wynn, T. and McGrew, W. C. 1989. An ape's view of the Oldowan. *Man* 24, 383-398.

Wynn, T., Hernandez-Aguilar, A., Marchant, L. F. & McGrew, W. C. 2011. "An ape's view of the Oldowan" revisited. *Evolutionary Anthropology* 20, 181–197.

Youngblade, L. M., and Dunn, J., 1995. Individual differences in young children's pretend play with mother and sibling: links to relationships and understanding of other people's feelings and beliefs. *Child Development* 66, 1472-1492.

Zentall, T. R. 2004. Action imitation in birds. *Animal Learning and Behavior* 32, 15-23.

Ziv, M., Solomon, A., Strauss, S., and Frye, D. 2016. Relations between the development of teaching and theory of mind in early childhood. *Journal of Cognition and Development* 17, 264-284.

Appendix 1

Participant Information Sheet (Face to Face)

Study Title: Identifying Cognitive and Linguistic Ability in the Archaeological Record: A Knapping Training Study Exploring Emulation, Imitation, and Teaching

Researcher: Cory Cuthbertson **Ethics number**: 17375

Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form.

What is the research about?

How can stone tools tell us about how language evolved? This study is part of a PhD project at the University of Southampton looking at how people learn to make stone tools when taught in different ways. Different types of teaching require different mental abilities; if we can measure the tools to find differences between groups that used different mental skills (for example, did some groups make tools that were much more similar to each other because of the way they were taught?), researchers might be able to detect which types of teaching were being used in ancient communities. From there, it would support the presence of certain mental capabilities, including language abilities, of the makers of ancient stone tools.

Why have I been chosen?

This study is looking for participants that consider themselves fit and healthy, consider themselves to have good vision (with or without glasses), and are between the ages of 18-65.

What will happen to me if I take part?

If you choose to take part in this study, you will learn how to make stone age handaxes, in a class with five others, over a series of three lessons. It does not take a lot of strength to knap a handaxe, but is more about precision and planning where to hit – a bit like snooker.

You will be asked to come to three training sessions, each lasting 2 hours, where you will learn stone knapping. The sessions will be indoors, at Fusion Arts Centre on Cowley road, near many bus connections.

You will either be taught by an instructor that interacts normally with you, a silent instructor, a video to copy from, or from simply tools in front of you that the instructor made previously. You won't have a choice in what group you are placed in. You will be required to wear protective eyewear and gloves. This will minimize the risk of getting chips of stone in your eyes, or cuts on your body. You will need to have your legs and feet covered (no open toed shoes or shorts; jeans are advised), and to wear clothing you don't mind getting a bit dusty.

When you first come, you will have a chance to meet the instructor and the others in your group, and the session will be explained to you. You will be asked to sign a consent form if you agree to take part in the study. You can, however, stop at any time. There will also be a break half way through the lesson. You will be asked to take a seat in a chair and you will have a cardboard box on the floor to collect your knapping debris. The instructor will then show you how to knap a piece of porcelain (about 700g) by hitting it with a quartzite hammerstone (about 350g). The porcelain blanks have been moulded into the same shape, so that everyone starts with exactly the same shape of piece. When you are finished, you tell the instructor who will take your finished tool and box, and give you another box and blank to continue practising with.

If you would like to hear about how this research progresses, the researcher will collect your email address and send you the link to the study when it is published as a PhD thesis 2016.

Are there any benefits in my taking part?

Your involvement will help me to complete this experiment which generates the data for my PhD project. It will aid in adding knowledge to the experimental archaeology discipline. It should also be a fun skill to learn, and may encourage you to take up a challenging and enjoyable hobby.

Are there any risks involved?

Safety gear should minimize the risks, but there is a chance you may cut yourself or damage your clothes on the pieces of chipped stone. Please always wear the safety gear if you or any other participants are knapping. There is a risk you may hit your fingers or your leg with a hammerstone, or drop something on your feet. Please wear durable clothing you do not mind getting a bit dusty, and closed toe shoes to protect your feet.

Will my participation be confidential?

The data collected will be confidential and kept on a password protected computer. Your name will only be available to the researcher, and will not be published alongside the collected data.

What happens if I change my mind?

You have the right to withdraw from the study at any time without your legal rights being affected.

What happens if something goes wrong?

If you have a concern or complaint, you should contact the Chair of the Faculty Ethics Committee Prof Chris Janaway (023 80593424, c.janaway@soton.ac.uk).

Where can I get more information?

If you would like any other information or have any questions, please contact Cory Cuthbertson at c.m.cuthbertson@soton.ac.uk, or on 07854649870.

Appendix 2

CONSENT FORM *(FACE TO FACE: 1.0)*

Study title: Identifying Cognitive and Linguistic Ability in the Archaeological Record: A Knapping Training Study Exploring Emulation, Imitation, and Teaching

Researcher name: Cory Cuthbertson

Staff/Student number: 26423219

ERGO reference number: 17375

Please initial the box(es) if you agree with the statement(s):

I have read and understood the information sheet and have had the opportunity to ask questions about the study.	
I agree to take part in this research project and agree for my data to be used for the purpose of this study	
I understand my participation is voluntary and I may withdraw at any time without my legal rights being affected	

Data Protection

I understand that information collected about me during my participation in this study will be stored on a password protected computer and that this information will only be used for the purpose of this study. All files containing any personal data will be made anonymous.

Name of participant (print name)..

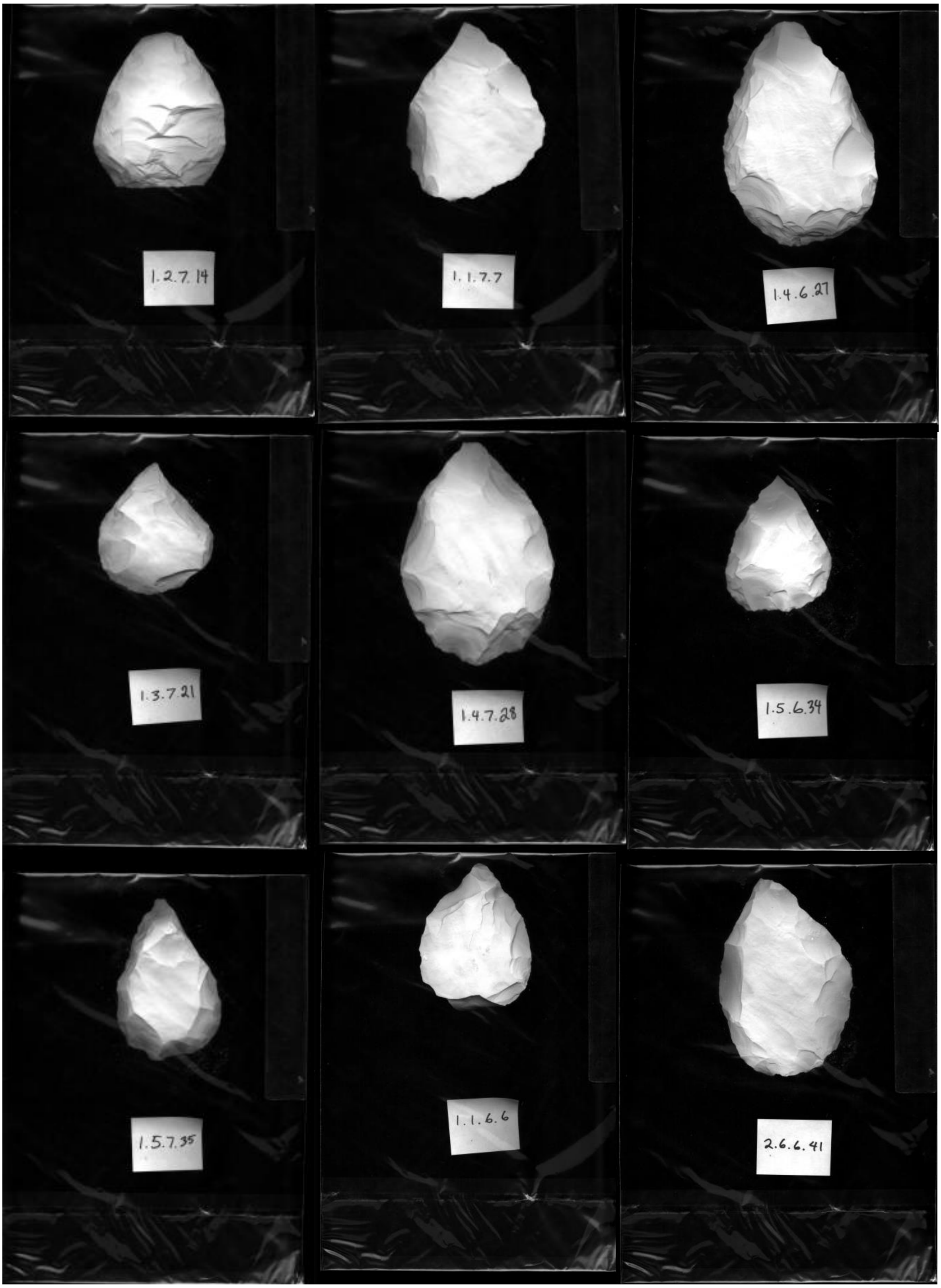
Signature of participant..

Date..

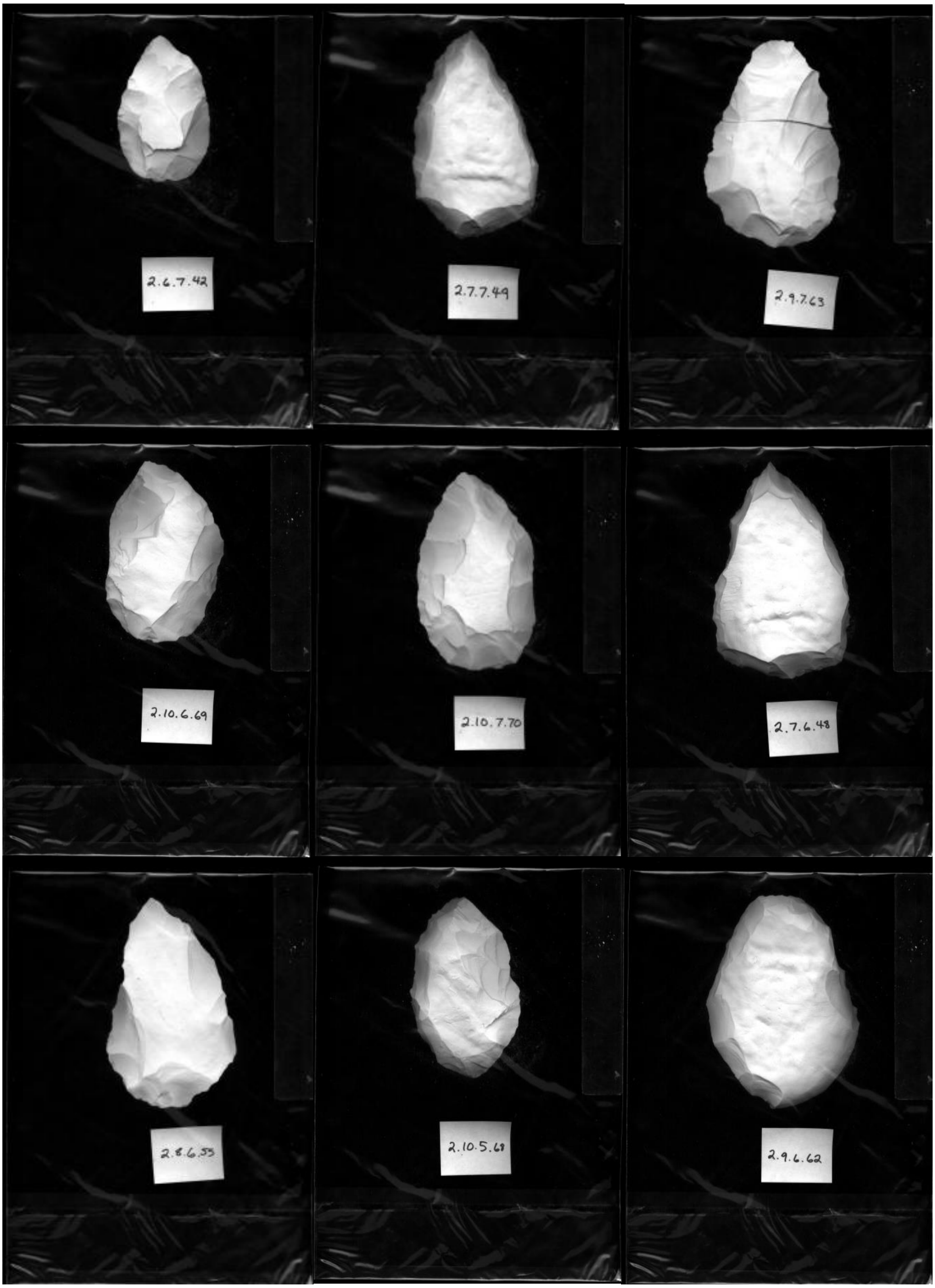
Appendix 3

Select Photos of Experimental Handaxes

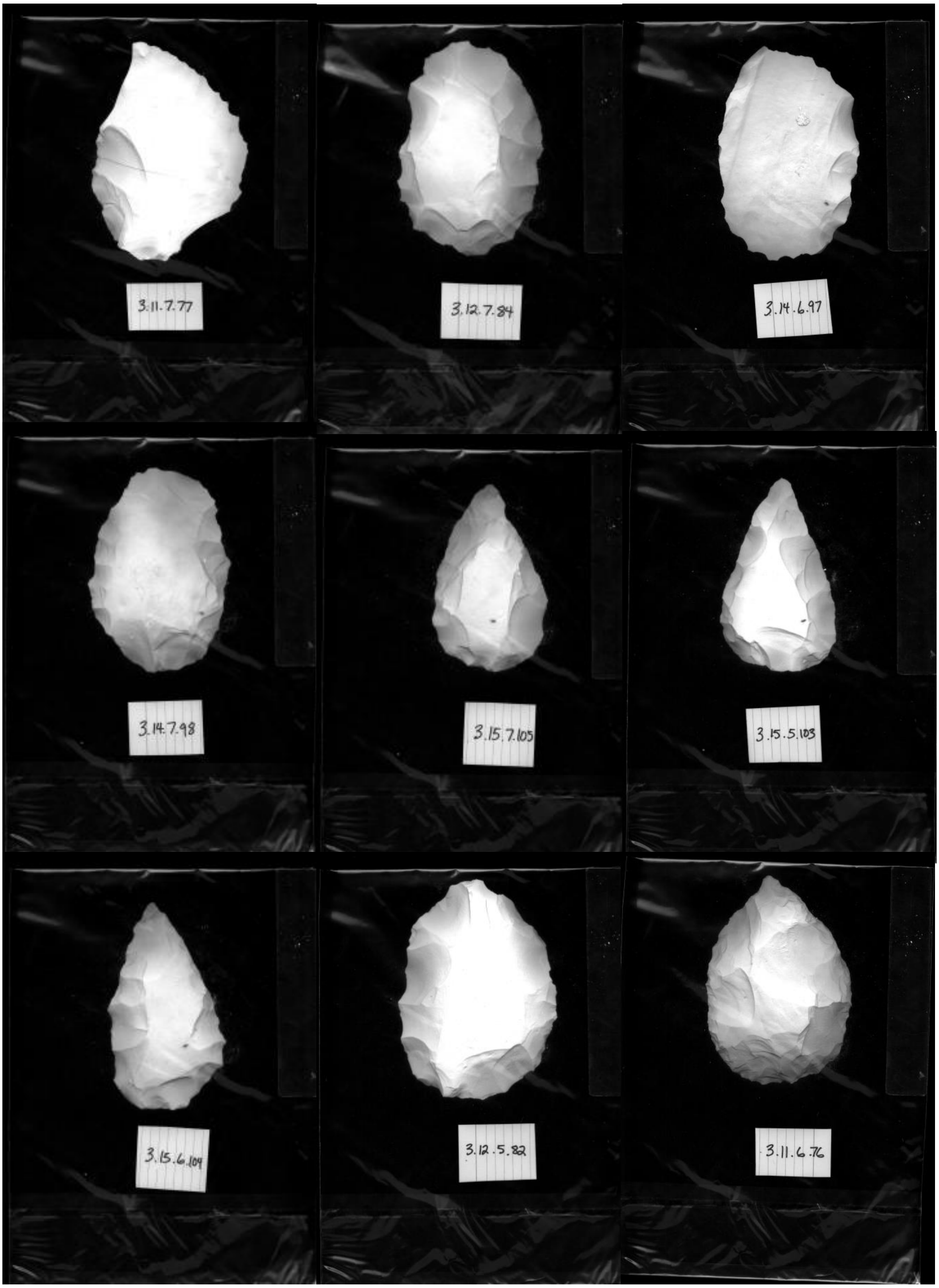
Group 1 (Emulation)



Group 2 (Imitation)



Group 3 (Silent)



Group 4 (Verbal)