Publicly Accessible Penn Dissertations

1-1-2014

Experimentation and Scientific Inference Building in the Study of Hominin Behavior through Stone Artifact Archaeology

Sam Chieh-Heng Lin
University of Pennsylvania, samlin@sas.upenn.edu

Experimentation and Scientific Inference Building in the Study of Hominin Behavior through Stone Artifact Archaeology

Abstract

Since the beginning of prehistoric archaeology, various methods and approaches have been developed to describe and explain stone artifact variability. However, noticeably less attention has been paid to the ontological nature of stone artifacts and the adequateness of the inferential reasoning for drawing archaeological interpretations from these artifacts. This dissertation takes a scientific perspective to rethink critically the ways that current lithic approaches generate knowledge about past hominin behavior from stone artifacts through experimentation (Chapter 2), and further, to explore the use of controlled experiments and uniformitarian principles for deriving inferences. The latter is presented as two case studies about Late Pleistocene Neanderthal behavior in southwestern France (Chapter 3 & 4).

Archaeological reasoning is inescapably analogical, and archaeological knowledge is bound to be established on the basis on modern observations. However, simplistic treatments of archaeological analogs often result in inferences of questionable validity. In this dissertation, it is argued that greater attention is required to consider the implication of experimental design, variable control, and analogic reasoning in the construction of archaeological inference from stone artifacts. It is argued that the ability to move beyond the constraint of modern analogs in archaeological knowledge production lies in the use of uniformitarian principles that operate independently from the research questions archaeologists wish to evaluate.

By examining the uniformitarian connection between platform attributes and flake morphology, the first case study explores how the production of unretouched flakes can be altered in ways that increase their relative utility, as reflected in the ratio of edge length to mass. Application of this relationship to Middle Paleolithic assemblages shows two modes of flake production pattern, possibly related to different ways Neanderthal groups managed the utility of transported tool-kits. The second case study applies a geometric model to assess the lithic cortex proportion in the Middle Paleolithic study assemblages. An excess or deficit of cortex relative to artifact volume provides an indication of possible artifact transport to or from the assemblage locality. Results show correlation between assemblage cortex proportions and paleoenvironmental conditions, suggesting possible shifts in Neanderthal artifact transport pattern and land use during the late Pleistocene.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Anthropology

First Advisor
Harold L. Dibble

EXPERIMENTATION AND SCIENTIFIC INFERENCE BUILDING IN THE STUDY

OF HOMININ BEHAVIOR THROUGH STONE ARTIFACT ARCHAEOLOGY

Sam Chieh-Heng Lin

A DISSERTATION

in

Anthropology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2014

Supervisor of Dissertation

Signature _____

Harold L. Dibble, Professor, Anthropology

Graduate Group Chairperson

Signature _____

Clark L. Erickson, Professor, Anthropology

Dissertation Committee:

Shannon P. McPherron, Researcher, Max Planck Institute of Evolutionary Anthropology

David R. Braun, Assistant Professor, The George Washington University

Deborah I. Olszewski, Adjunct Associate Professor, Anthropology

Theodore G. Schurr, Professor, Anthropology

EXPERIMENTATION AND SCIENTIFIC INFERENCE BUILDING IN THE STUDY

OF HOMININ BEHAVIOR THROUGH STONE ARTIFACT ARCHAEOLOGY

ACKNOWLEDGMENT

ABSTRACT

EXPERIMENTATION AND SCIENTIFIC INFERENCE BUILDING IN THE STUDY

OF HOMININ BEHAVIOR THROUGH STONE ARTIFACT ARCHAEOLOGY

Sam C. Lin

Harold L. Dibble

Since the beginning of prehistoric archaeology, various methods and approaches

have been developed to describe and explain stone artifact variability. However,

noticeably less attention has been paid to the ontological nature of stone artifacts and the

adequateness of the inferential reasoning for drawing archaeological interpretations from

these artifacts. This dissertation takes a scientific perspective to rethink critically the

ways that current lithic approaches generate knowledge about past hominin behavior

from stone artifacts through experimentation (Chapter 2), and further, to explore the use

of controlled experiments and uniformitarian principles for deriving inferences. The latter

is presented as two case studies about Late Pleistocene Neanderthal behavior in

southwestern France (Chapter 3 & 4).

Archaeological reasoning is inescapably analogical, and archaeological

knowledge is bound to be established on the basis on modern observations. However,

simplistic treatments of archaeological analogs often result in inferences of questionable

validity. In this dissertation, it is argued that greater attention is required to consider the

implication of experimental design, variable control, and analogic reasoning in the

construction of archaeological inference from stone artifacts. It is argued that the ability to move beyond the constraint of modern analogs in archaeological knowledge production lies in the use of uniformitarian principles that operate independently from the research questions archaeologists wish to evaluate.

By examining the uniformitarian connection between platform attributes and flake morphology, the first case study explores how the production of unretouched flakes can be altered in ways that increase their relative utility, as reflected in the ratio of edge length to mass. Application of this relationship to Middle Paleolithic assemblages shows two modes of flake production pattern, possibly related to different ways Neanderthal groups managed the utility of transported tool-kits. The second case study applies a geometric model to assess the lithic cortex proportion in the Middle Paleolithic study assemblages. An excess or deficit of cortex relative to artifact volume provides an indication of possible artifact transport to or from the assemblage locality. Results show correlation between assemblage cortex proportions and paleoenvironmental conditions, suggesting possible shifts in Neanderthal artifact transport pattern and land use during the late Pleistocene.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 1: Introduction

1.1 Stone Artifacts and Archaeological Knowledge of the Past

The goal of anthropology can be said to analyze and explain the total range of physical and cultural similarities and differences that are characteristic of the entire spatial-temporal span of human existence. For the study of human evolution and the emergence of our species, archaeology plays a pivotal role in supplying the basic Paleolithic framework that remains fundamental to much of the paleoanthropological discussion taking place today. From the Oldowan, Acheulian, and Mousterian, to the various Upper Paleolithic industries, the transformation of stone artifacts through time is seen to reflect evolution in hominin behavior and technological capabilities associated with biological evolution (Ambrose 2001; Foley and Lahr 2003).

During the 1970s and 1980s, the growing interest of anthropological archaeology in North America (Binford 1962, 1965; Longacre 1970; see Trigger 2006) led research in the subsequent decades to focus on the organizational dynamics and anthropological processes underlying the formation of stone artifact assemblage variability, including mobility pattern, technological design, and subsistence strategies; behavioral processes that likely implicate hominin adaptation and evolution (e.g., Bamforth 1986, 1990; Binford 1977, 1979, 1980; Nelson 1991; Torrence 1983, 1989). More recent emphasis on the connection between material culture and the socio-demographic dimensions of behavior offers novel archaeological perspective to past hominin behavior (e.g., Shennan

2008, 2011). Because stone artifact production involves technological knowledge and skills that are necessarily transmitted through learning from one individual to another, the continuity and change in artifact variability over time signals historically-derived processes that are contingent on wider social, demographic, and biological factors (e.g., forms of learning, social network and demographic structure, and cognitive capability such as long-term memory and information storage) (Henshilwood and Dubreuil 2011; Lycett and Norton 2010; Lycett 2009b; Lyman and O'Brien 1998; O'Brien and Holland 1990; O'Brien et al. 2001, 2010; Powell et al. 2009; Shennan 2011, 2000, 2008; Tennie et al. 2009).

The development of these frameworks is crucial as they allow archaeological inferences to move beyond the objects themselves to a higher level of evolutionarily interesting issues, and from there to contribute to wider multidisciplinary discussions of human evolution. For example, the connection between material culture, neurological processes, and other cognitive aspects, including linguistic ability, has been repeatedly emphasized (e.g., Dediu and Levinson 2013; Stout and Chaminade 2009; Stout 2011; Stout et al. 2008, 2011; Wynn 1995, 2008; see McPherron 2013 for review). As such, interpretations of lithic technological complexity are now commonly featured as supporting evidence for past hominin brain evolution in terms of cognitive, symbolic, and language capacities (e.g., d'Errico et al. 2003; Pelegrin 2009; Stout 2011; Toth et al. 1993; Wynn and Coolidge 2004; Wynn 2008; c.f. Dibble 1989). Likewise, archaeological data are increasingly sought by researchers in fields such as paleogenomics and physical anthropology to validate and contextualize hominin population models with cultural,

behavioral, and demographic details about the various hominin groups in question (e.g., Lalueza-Fox and Gilbert 2011; Lalueza-Fox et al. 2011; Pearce et al. 2013).

For the study of stone artifacts, the wide array of archaeological knowledge generated in relation to past human behavior and evolution can be separated into four main domains: typology, technology, function, and cognition (following the classification proposed by Haidel 2007). While these domains are not mutually exclusive, they provide the main analytical structure for examining the behavioral context of past hominin groups in terms of subsistence, settlement, and social organization as well as the continuity and change in cultural behavior and technical innovation. Each of these domains is briefly summarized here.

1.11 Typology

The typological domain involves the characterization of artifacts and artifact assemblages based on artifact form and style. Specific stone artifact types that recur in assemblages over time are viewed as typifying prehistoric industries, populations, or groups (Haidel 2007). Definition of these archaeological entities is often based on particular artifact type(s) and/or their relative frequency within an artifact assemblage. The distribution of these industries across time and space, in turn, provides the culture-historic framework of the prehistoric world that is commonly used today in Paleolithic studies. The more recent cultural transmission theory provided a supporting basis for this approach (Boyd and Richardson 1985; Dunnell 1980, 1989; Lycett 2010, 2011; O'Brien and Holland 1990; Shennan 2008, 2011). In this view, the continuity of artifact form is

3

seen to reflect processes of knowledge transmission where a particular style represented in material culture was selectively passed on. Based on the concept of formal similarities and heritable continuity, the persistent presence of specific artifact types or attributes through time signals positive or stabilizing selective process where the learning and practice of particular technical knowledge was maintained or encouraged (Collard and Shennan 2008; Lyman and O'Brien 1998; Shennan 2008). The degree of variation reflects constraints on learning and the levels of conformity (Shennan 2008). This historical connection in material culture allows archaeologists to reconstruct phylogenetic relationships between artifact forms through time and space (Clarkson 2010; Clarkson et al. 2012; Lycett and von Cramon-Taubadel 2008, forthcoming; Lycett 2007, 2008, 2009a,b, 2010; Lycett et al. 2010), which are then taken to represent the culture-history of prehistoric populations.

1.12 Technology

The technological domain of archaeology offers a view of the broader behavioral sequence in the production and use of stone artifacts. It focuses on the role of raw material procurement, the operational sequence of core reduction, the reduction intensity and life history of artifacts, and artifact selection and transport (Bleed 2001). These technological activities are often considered within the wider context of land use pattern and foraging strategy. Much attention focuses on understanding the dynamic behavioral processes that underlie the formation of lithic assemblages in relation to the associated economic and environmental conditions. This sort of study examines the organization of technology in relation to mobility patterns, raw material economy, artifact function,

4

design, and maintenance, and risk management (e.g., Blades 2003; Delagnes and Rendu 2011; Dibble and Rolland 1992; Feblot-Augustins 1993; Fernandes et al. 2008; Rolland and Dibble 1990; Roth and Dibble 1998; Wallace and Shea 2006). More recent emphasis on the *chaîne opératoire* concept stresses the importance of the overall technological system represented in lithic assemblages. Here, the goal is to capture the sequence of mental operations and technical gestures in prehistoric stone artifact production (Bar-Yosef and Van Peer 2009; Pelegrin 1990; Perlés 1992; Sellet 1993). Because each stage of the technical sequence involves options influenced by technical, economic, social, and cultural factors, the recurrent combinations of these sequences of operation – sometimes referred to as "strategies" – are seen as technical traditions, or 'knowledge', shared among group members (Dobres 2000; Perlés 1992).

In many ways, the technological domain provides a more coherent perspective on cultural transmission. Instead of focusing solely on the morphology of particular artifact type as the marker of technical knowledge, the emphasis on sequence implies that the transmission of knowledge encompasses the entire technical system. It involves not only the production of specific artifact types but also various other aspects of technological behavior, including raw material selection, core volume organization, artifact transport, and mobility pattern (Boëda et al. 1990; Delagnes and Meignen 2006; Delagnes and Rendu 2011; Delagnes 2010; Delagnes et al. 2007; Geneste 1985, 1991; Meignen et al. 2009; Perreault et al. 2013). Indeed, much of the recent characterization of Middle Paleolithic variability has focused on technical sequences as oppose to the traditional assignment of Mousterian type facies (e.g., Delagnes and Meignen 2006; Delagnes and

Rendu 2011; Faivre et al. forthcoming). Furthermore, comparison between sequences

also sheds light on the differences and similarities between industries in terms of

manufacturing steps and end-product designs in relation to broader dynamics of

population interaction, information exchange, and acculturation (e.g., Kuhn and Zwyns

2014; Roussel 2013; Tostevin 2013; Zwyns 2012).

1.13 Function

The functional domain of archaeology refers to the study of artifact function and

use. A greater interest in artifact function emerged in the 1970s and 1980s when the

discussion over the nature of stone artifact variability led various research efforts to

explore more objective and definite means for assessing artifact function (e.g., Brink

1978; Briuer 1976; Hayden 1979; Semenov 1964). Today, microwear and residue studies

are commonly featured in Paleolithic archaeology. These kinds of studies provide

information on the type of use and activity for which artifacts were employed, and also

provide a way to examine changes in artifact functionality associated with the appearance

of new technologies. For example, experimental comparisons of the morphology and

distribution of fractures on flakes have pointed to possible origins of projectile

technology in the Middle Paleolithic and the Middle Stone Age (Lombard and Pargeter

2008; Sahle et al. 2013; Shea 1987, 1988, 2006; Sisk and Shea 2011; Villa and Lenoir

2006; Wilkins et al. 2012, 2014; Yaroshevich et al. 2013; but see Iovita et al. 2014;

Pargeter 2011; Rots and Plisson 2014). Furthermore, as Haidel (2007) pointed out,

studies of the functional domain also give evidence about technologies that left little

direct archaeological trace, such as the manipulation of organic raw materials (e.g., Soffer

2004). This aspect is particularly important as stone artifacts may not have constituted the main component of past technological complexes. Instead, many activities were likely carried out with implements made from organic materials, as evidenced by, for example, the wooden spears from Schöningen (Thieme 1997). The identification of resin residue on artifacts possibly used as binding agents for hafting also provide evidence for the way stone artifacts were used in the past (Boëda et al. 1996, 2008; Lombard 2004, 2005, 2006, 2007, 2008; Wadley 2005, 2010; Wadley et al. 2004).

1.14 Cognition

The cognitive domain concerns the neuro-cognitive background for hominin behaviors. In a sense, this is a higher level inference through combining information derived from the three previous domains to further investigate the range of behavioral choices and decision-making carried out by past hominin groups in relation to particular environmental conditions and social settings. What this domain offers is the ability for archaeologists to interpret the socio-cultural processes and cognitive capabilities of Paleolithic populations (Haidel 2007). For stone artifacts, the production of flakes requires the knapper to successfully articulate and combine various motor actions with respect to the physical properties of stone fracture as well as the geometric configuration of the stone surfaces (Moore 2011; Stout and Chaminade 2009; Stout 2011).

While it is true that the cognitive capabilities underlying tool making, including forms of pattern recognition, rule abstraction, motor coordination, associative learning, and understanding of material properties, have been demonstrated to exist in extant

primates and presumably early hominins (Haslam et al. 2009; Matsuzawa 2001; Schick et al. 1999; Toth et al. 1993), consciousness is likely essential when the sensory pattern that controls motor coordination is extended spatially and temporally (Rossano 2009). In other words, artifact forms and manufacture sequence that appear to exhibit some form of standardization and a minimum level of expertise reflect not only the simple pursuit of obtaining a cutting edge but also a conscious control of multiple aspects of the technology, which signal the active involvement of long term memory, abstraction, and practice (Rossano 2003; Wynn 1981, 2002). The increased consciousness behind tool making is also viewed as a critical factor that enhanced the cognitive fluidity and creativity of hominins (Rossano 2009). Because technological advancement and creativity is highly integrated and is determined by cognitive and neurological structures, it has been argued that insights into prehistoric cognitive capacity may be gleaned from the technicality of artifact production. For example, based on the model of design space and the framework of technological innovation and cognition from psychology, Moore (2007) argued that the lack of hierarchically combined technological "units" in the lithic assemblages associated with *Homo floresiensis* may indicate a lack of cognitive capacity in creativity and technological inventiveness (also see Moore and Brumm 2009).

1.2 Some Basic Research Questions in Stone Artifact Archaeology

These four domains capture the majority of current research effort in stone artifact archaeology within the context of paleoanthropology. In particular, typological classification of archaeological industries provides the chronological structure for the basic divisions upon which behavioral and cognitive interpretations are made.

Assemblages of archaeological finds recovered from sites represent material remnants resulted from particular sets of behavior from specific groups of people at a certain point in time. The ability to attribute artifact and assemblage type to certain 'people' means archaeological variability can be seen to reflect differences in the nature of these groupings (populations, cultures, or biologically distinct hominin groups).

While these frameworks have been productive in extending archaeological perspectives to paleoanthropological discussions, it is important to recognize that there exist a number of basic ontological and epistemological questions in these frameworks regarding the inferential linkage between stone artifact categories and higher level interpretations of hominin behavior and adaptation. These questions include: What is the nature of stone artifacts and how do they relate to human behavior? How do the artifact categories identified by archaeologists relate to the actual intentionality involved in the production and use of these artifacts in the past? Such theoretical issues are central to the current Paleolithic archaeology research agenda and impact many of the ways inferences are drawn from the lithic archaeological record. Yet, as researchers progressively move towards addressing wider evolutionary topics, discussions over the way that these research questions implicate the integrity and confidence of the resulting inference remain less apparent in the current literature (Shea 2011). As archaeological interpretations become increasingly integrated with broader paleoanthropological frameworks, these issues should be thoroughly considered.

The root of some of these inferential problems can be attributed, in part, to the fundamental difference between stone artifact archaeology and other fields of paleoanthropology. In physical anthropology, zooarchaeology, and paleogenomics, the connection between the subject of study and the actual organism of interest (hominins or primates) is relatively concrete. When a hominin fossil is uncovered, one can be certain that the bone came from a specific individual at a certain point in time. In other words, the connection between the bone and the biological reality of the hominin is clear. While, in most cases, we can be reasonably confident that the stone artifacts recovered by archaeologists were created by past hominins (but see Chase et al. 2009; Dibble et al. 2006), it is more difficult to immediately assign further behavioral or biological reality to these implements (discussed more below). Because of this disjuncture between stone artifact and past "people", much of the inferential logic in archaeological interpretations concerning lithics focuses on specific sets of epistemological connections between artifacts and broader behavioral or biological phenomena. In particular, classificatory units generated by archaeologists (etic categories) are utilized to capture aspects of reality concerning the intention of past people (emic categories). This theoretical distinction is not new and has played a central role in major archaeological discussions, such as the Ford/Spaulding debates in the 1950s (Ford 1954a–c; Spaulding 1953, 1954). Yet, they remain largely unresolved and continue to carry significant metaphysical consequences in the interpretations of hominin behavior from stone artifacts.

1.21 Etic vs. Emic – The Nature of Artifacts, Manufacture Sequence, and Intentionality

Because of the inability to attribute prehistoric stone artifacts to any historically-known phenomena, early Paleolithic archaeologists adopted classificatory schemes from the natural sciences, particularly geology and paleontology, with the goal of organizing the observed variability into meaningful units (Van Riper 1993; Rolland and Dibble 1990). Specific artifacts were recognized for their recurring forms and/or secondary retouch, which led them to represent categories of intentional design and production (Davidson and Noble 1993; Davidson 1991, 2002). The debate between François Bordes and Lewis Binford in the 1960s about the functional and stylistic nature of artifacts signaled the wide-spread perception that artifact types – as created by archaeologists – held behavioral and cognitive reality about the past hominins who manufactured and used these objects. In many ways, this construct is deeply embedded in archaeological thought and can be seen in the terminology of lithic studies. The common referral of stone artifacts as "tools" reflects the assumed connection between lithic materials and their intended use or design. Also, because these tools represent discrete and mutually exclusive types, they are taken to correspond to particular designated functions such as cutting, scraping, or chopping (e.g., Schoville 2010). This assumed correlation is particularly apparent in the discussion of pointed artifacts and projectile weaponry (e.g., Brookes et al. 2006; Wilkins et al. 2012).

What this discussion illustrates is the archaeologists' goal of understanding the intention of past people through material culture. Indeed, the argument between James

Ford and Robert Spaulding in the 1950s demonstrated how differences in the conception of analytical categories carry metaphysical implications about archaeological interpretations of intentionality (O'Brien and Lyman 2002). While Ford (e.g., 1954) argued that classification of artifact types only serve as analytical units that help archaeologists interpret the grouping of assemblages (i.e., they are constructed by the analysts and are in no way 'real' to the people in the past), Spaulding (1953:305) held that the goal of artifact classification is the discovery "of combinations of attributes favored by the makers of the artifacts, not an arbitrary procedure of the classifier". From this view, real types are inherent in artifacts, and, by collecting and analyzing enough data, these 'emic' categories can be inductively derived. Statistical techniques such as discriminant analysis were seen to provide objective means for determining these types by assessing which attributes co-occur at frequencies greater than that allowed by chance alone (Spaulding 1953). In other words, recurring combinations of attributes demonstrate intentional choice and decision in the manufacture of artifacts (O'Brien and Lyman 2002).

In a largely implicit way, the emic perspective characterizes much of the use of artifact types today in Paleolithic archaeology. As discussed before, because retouched and formal artifacts possess morphological characteristics that suggest intentional modification and design as opposed to variation allowable by chance, these types are considered to reflect prehistoric cultural norms or group-specific practice. Historically, the interpretation of past peoples has been centered on these artifact types, such as bifaces, Levallois end-products, scrapers, points, blades, and microlithic artifacts. Since

these 'tools' are identified as intentionally shaped end objects, their manufacturing

sequence from the initial raw material can be reconstructed, and hence "debitage", "by-

products", or "waste" can be readily identified from the assemblage.

A significant part of lithic studies have been dedicated to the study of the

manufacturing sequences of these tool types (e.g., Boëda 1988, 1993, 1994, 1995;

Bourguignon 1997; Forestier 1993; Magne and Pokotylo 1981). This emphasis is most

strongly demonstrated by studies that view technology as holistic chains of events

governed by individual- or group-specific mental templates or knowledge (Dobres and

Hoffman 1994; Dobres 2000; Perlés 1992). However, the search for intentionality in

stone artifacts through the reconstruction of production sequence leads to three issues.

First, because the operational sequence of lithic reduction is perceived as sequential and

linear, the chain of technical operation has to be driven by a pre-existing goal. This goal

is often framed as the production of certain desired end-product at the end of the

operational chain (Dibble & Bar-Yosef 1995). As Bar-Yosef and Van Peer (2009)

pointed out, the identification of archaeological end products is a modern construct based

on our analytical framework and cannot be extended to an emic designation of how these

items were perceived in the past. It is true that certain artifact types, such as projectile

points and adzes, were indeed rigidly produced for designated functions, and their

morphology clearly associates with designs for enhancing efficiency and functionality.

Then again, in most cases, these formal artifacts compose a small fraction of the overall

assemblage and the life history of these artifacts may well span over multiple generations

of individuals. In other words, without making *a priori* assumption that certain artifact

types are in essence 'real' to past technological systems, it is impossible to discern what the production chain and associated by-product or waste actually was.

The other issue is that the linear sequence of these operational chains requires the underlying mental template to be a singular flow carried out by an individual or several individuals sharing similar mental template over the reduction sequence (Bar-Yosef and Van Peer 2009). In other words, in order for reduction sequences to be reconstructed from an assemblage, it must be assumed that individuals contributing to the formation of an assemblage all conformed to an identical standard of lithic production that characterizes the assemblage (Bar-Yosef and Van Peer 2009). This assumption is difficult to sustain given the coarse resolution of the Paleolithic record, not to mention the near impossibility of establishing that an assemblage or any of its sub-divisions correspond to contemporaneous activities or even ones that occurred over a short time interval.

For this reason, studies have increasingly employed refitting as a way of controlling chronology and to find high-resolution events within the archaeological record (Chiotti et al. 2007; Close 2000; Vaquero 2008; Vaquero et al. 2012). While refittings do indicate sequential events in the reduction process, the way that these events relate temporally cannot be assumed as they could easily be attributable to unrelated actions by individuals separated by considerable amounts of time. Furthermore, as refitable elements most often only represent a subset of the assemblage, the connection of temporality and intent between these artifacts and the rest of the assemblage is difficult to establish.

The third issue relates to the conception that lithic assemblages from sites represent complete operational sequences involving coherent sets of actions and decisions. This perspective contrasts with ethnographic records which tend to suggest flexibility and fluidity in the production and use of stone artifacts, where the desired artifact or attribute differ by context or individual (Holdaway and Douglass 2012; Moore 2003). In other words, a seemingly sequential process of lithic reduction may involve multiple individuals with varying intentions and views that are largely unrelated from each other. As Moore (2011) demonstrated, a sequence of flake production or the creation of recurring artifact forms does not necessarily require the involvement of higher-level intention and sequential planning. Instead, these patterns can be created by repeated actions of flake removal with basic recognition of core geometry.

Furthermore, while the production of every flake clearly involves some form of intention, determining how these intentions relate to those that govern the selection of usable flakes is a different matter. As Hiscock (2004; also see Holdaway and Douglass 2012; Turq et al. 2013) demonstrated in his ethnographic study, the process of flake production and selection may be performed at different stages and by different individuals with varying selection criteria. An archaeological reduction sequence as identified by archaeologists therefore could result from multiple unrelated 'sequences' involving many intentions concerning production and use. As a consequence, an ethnographic 'type' recognized emically could vary considerably in its forms between individuals, even within groups that share cultural identities and socially-conscious groupings (see White and Thomas 1972). These observations of stone artifact use in a

living context call into question the notion that archaeologists are able to distinguish complete technological sequences in lithic assemblages.

1.3 Stone Artifacts and Evolution

These issues illustrate the research challenges in archaeology of describing stone artifacts and other archaeological phenomenon as meaningful units that can be further related to aspects of hominin lifeways and evolution. They also demonstrate the aspiration shared by archaeologists to address evolutionary interesting questions. If the goal of stone artifact archaeology is to understand human evolution, then the critical question lies in delineating the relationship between stone artifact variability and the evolutionary fitness of the hominin toolmakers/users. Ultimately, of course, it is the organism that is the unit of natural selection and not the stone artifacts, and most archaeologists would agree with this statement. Nonetheless, demonstrating the linkage between stone artifacts and evolutionary fitness remains largely problematic.

Discussions of this subject are often framed with a view that artifacts represent the technological medium for solving subsistence problems or achieving survival goals. In other words, artifacts are meaningful extensions of behavior and a proxy of selection on behalf of the tool users. Selection on artifacts hence is seen to have operated on the functionality of the tools for which they were designed and manufactured. Indeed, many studies have focused on examining the design of artifact forms and their relative effectiveness and economy for serving organizational needs and carrying out the designated tasks (e.g., Ahler and Geib 2000; Bleed 1986; Eren and Lycett 2012; Jennings

et al. 2010; Kelly and Todd 1988; Kelly 1988; Prasciunas 2007). The persistence or change in artifact form in turn reflects shifting technological or behavioral solutions in mediating people and the environment. Under a longer time scale, increased artifact complexity over time is seen to indicate innovation of more specialized and effective tools made possible by the greater cognitive and motor capability of hominin toolmakers.

This perspective is tantalizing and represents a theme commonly featured in stone artifact archaeology studies today (Ambrose 2001; Stout and Chaminade 2012; Stout 2011). However, if it is the tools that were selected and carry evolutionary significance in relation to the tool user, then two key assumptions concerning the relationship between tools and people must be satisfied. First, tools have to be an adaptive entity that fulfills specific functional purpose(s) regardless of the context of use – i.e., as an extended phenotype of the individual that possess an absolute functional or adaptive quality. Just like nests are always created to serve the particular function of incubating eggs and sheltering young birds, tools are designed to carry out specific functions. For some forms of tools, researchers can be relatively confident that tool function remained largely constant regardless of how they were used, e.g., projectile points. Thus, changes in the morphology of these points can be linked to the emergence and shifts in the overall projectile technology (e.g., spear vs. bow-and-arrow) and their relative efficiency in terms of factors such as velocity, aerodynamics, and penetrative power (Cheshier and Kelly 2006; Christenson 1986; Lipo et al. 2012; Shea and Sisk 2010; Shea 2006; Sisk and Shea 2009, 2011).

Then again, if we consider the design and actual usage of some modern tool types, this assumption may not always hold true, even for projectile points. For example, a hammer is designed for hammering nails through hard substrates. Yet, we know from experience that a hammer can and does, in fact, get used for various purposes under different circumstances. Likewise, the use of stone artifacts in the past probably also shifted widely and continuously. While we can say certain kinds of flakes, such as Levallois products, possess greater amounts of cutting edge and hence offer more utility for use (Eren and Lycett 2012; Kuhn 1994; Lin et al. 2013; Morrow 1996), it does not necessarily mean these flakes were exclusively manufactured and used for specific purposes. Instead, the use of Levallois flakes, as well as other flakes, likely changed depending on the tool user and where/when the task was performed. Similarly, both Oldowan chopper-chopping tools and Acheulian handaxes could be used as cutting/chopping tools, cores, and/or hammers by hominins at any given time depending on the task at hand.

Although it could be argued that these earlier tool forms represent multifunctional tools and were later replaced by more specialized tools, this conception reflects the second assumption – because designed "tool types" are fundamental to the way technology operates in the modern world, they are universal to all tool-using hominin populations. It is therefore possible to identify these types from other unintended materials of production. This is commonly done through refitting and replicative flintknapping experiments, although sometimes the distinction is based purely on morphological characteristics alone, particularly with retouched pieces that share

18

recurring forms. However, as illustrated in the previous section, stone artifacts do not

possess any quality in themselves in allowing archaeologists to differentiate what is

wanted from the unwanted on the ground of morphology and production sequence.

Specifically, there is no clear way for tying the production, use and discard of one object

to one individual within a short time frame. Instead, the complexity in the interaction

between human intention and the production and use of stone artifacts likely means that

the manufacture, use, discard, recycle, transport of one flake could stem from

independent events, perhaps separated by considerable amounts of time. In other words,

end-products represent an etic category based on modern research criteria and it is

difficult to justify them as emic types or units of selection in the past.

1.4 Towards an Archaeological Science of Stone Artifact Archaeology

The root to this issue of conflating etic and emic categories in stone artifact

archaeology can be attributed to the mismatch between 1) the ways archaeological units

are constructed from artifacts, versus 2) the ontological understanding of the ways the

stone artifact record was formed in relation to human behavior. The latter aspect concerns

the fundamental questions of where archaeological data come from, and how should

researchers actually go about interpreting these archaeological observations in

meaningful ways. Undoubtedly, many archaeologists would agree that the answer to the

first question is 'the archaeological record', which is composed of archaeological remains

and their contextual composition. It represents the empirical 'facts' created by past events

that have been preserved until today (Binford 1987). Through observation of the

archaeological record, archaeologists generate data from the body of empirical facts to make inferences about the past.

However, since observation is a modern phenomenon, it relies on the criteria that archaeologists deem profitable to observe (Binford 1987, 2001; Sullivan 1978). Then again, if we cannot observe the events that led to the formation of the archaeological record, how do we come up with analytical units that can allow us to meaningfully reconstruct these past events? In the 1960s and 1970s, the desire of many North American archaeologists to move archaeology towards an anthropological science (Binford 1962, 1965; Fritz and Plog 1970; Salmon 1975; Watson et al. 1971) meant a departure from the traditional empiricist's view, where data and interpretation are acquired through the 'natural work of the mind when freed from impediment' (Bacon, in Commins and Linscott 1947: 154). The central focus of this change is to establish concrete referential frameworks to connect the archaeological record to past behavioral processes.

1.41 Units of Measurement versus Units of Interpretation

Despite advances afforded by these early studies associated with the New Archaeology movement, the etic/emic issue continues to undermine the inferential integrity of stone artifact archaeology in the 21th century. This issue signals that the theoretical issue runs even deeper in the production of archaeological knowledge that one might think. One alternative way to approach this issue is to argue as follows: if classification units are etic categories constructed by archaeologists, then it is necessary

for these units to be constructed on the basis of principles that are independent from the interpretation that archaeologists seek to make. If we define stone artifact types based on etic criteria but treat them as emic in nature, subsequent interpretation would inevitably mirror the definitions we have constructed in the first place (tautological). As a result, explanation of variability can only be achieved by granting these classifications some sort of *ad hoc* behavioral or cognitive reality in the form of either cultural preference or functional design (Dibble and Rolland 1992).

If the goal of stone artifact archaeology is to understand the formation of archaeological variability in relation to human evolution, then such interpretations necessarily assume that the behavioral activities carried out by past hominins were based on individual decisions and intentions. Principles for classifying and organizing archaeological facts into data are therefore required to operate independently from assumptions concerning past intention. In other words, the *units of measurement* for deriving empirical data from the archaeological record must be separate from the *units of interpretation* upon which archaeological inferences are constructed. This rule may seem obvious to many fields of formation studies, including taphonomy and geoarchaeology, where their inquiry begins with observations based on processes that are uniformitarian and largely independent from human intentions.

For many studies in archaeology this rule may seem more difficult to follow. More specifically, it is not clear how we can understand human behavior if we only focus on invariant processes unrelated to human intention. In stone artifact archaeology,

flintknapping experiments serve as one of the major areas for deriving referential

knowledge for interpreting stone artifacts. Studies of this sort replicate the production or

use of specific artifact types as a way to generate behavioral, or sometimes cognitive,

analogs that can be projected to the past. However, the predicament presented here is that

in order to replicate the production sequence or use of certain artifacts, archaeologists

must first identify which artifacts are useful to replicate – i.e., the need to assume *a priori*

the artifact types people in the past wanted to make and use. As a consequence, units of

interpretation become confounded with units of measurement, and the final inference

reflects more of the archaeological unit's presumed significance rather than the nature of

the independent observation.

1.42 Archaeological Science and Uniformitarianism

A possible way to break out of this cycle of inference fallacy is to seek examples

of knowledge production in the archaeological sciences. Fields in archaeological science

apply scientific methods to the archaeological material. Here, what makes a method

'scientific' is its reliance on theories that were developed through repeated empirical

observations, rigorous analysis, and hypothesis testing *on the study subject* (Binford

2001). The invariant nature of these regularities means they likely operated in the same

way in the past. Thus, these theories can be used to interpret the possible causal factors

that contributed to the formation of the archaeological record as archaeologists study

today. For stone artifacts, the hominin activities and other behavioral processes that led to

their creation have long disappeared. Then again, stone artifacts are actual physical

materials they operate on a set of uniformitarian rules, including fracture mechanics, solid

geometry, geochemistry, and mechanics of materials. This means that archaeologists can confidently utilize etic units of measurement without imposing *a priori* interpretations or classifications.

Archaeological units, therefore, need to be constructed upon principles that are not only independent from interpretation but also uniformitarian in nature. It is critical that the framework utilized for determining which attributes are useful to record can be confidently assumed to have also operated in the distant past. What this requires is for archaeologists to ask very basic questions regarding the properties of the archaeological record. For example, how do stones break? What variables are important in governing the fracture of stones and what are their observable effects on stone artifacts? What variables affect the number and size of flakes produced from a nodule with a given set of attributes? How does artifact movement and transport affect assemblage composition? How do the various morphological attributes of flakes affect use in different tasks and how do they translate to observable wear patterns?

These are basic ontological research questions concerning the formation of the stone artifact record that requires thorough investigation. They do not require assumptions about the behavioral significance of the analytical units, and therefore provide concrete connections between observable pattern and dynamic processes that are proxies to past behaviors. Through controlled experiments (defined in the following chapter) and other approaches such as simulation and modeling, it is possible to isolate and delineate the effect of specific variables and, from there, to show how different

behavioral processes can alter the structure and distribution of these variables within the archaeological sample. In so doing, it is possible to falsify hypotheses and establish baseline patterns to which archaeologists can compare archaeological data. These sort of inference constructions, in turn, allow archaeological interpretations to be drawn in a more scientifically sound manner.

This perspective may give the impression that stone artifacts can tell us only mundane and trivial things about the human past. Higher exterior platform angle and platform depth result in larger flakes. Flakes from early reduction stages on average are bigger and contain more cortex. While one could attempt to dismiss these approaches for producing uninteresting "Mickey Mouse Laws" (Flannery 1973), these seemingly ordinary principles, coupled with the ubiquity of stone artifacts, are in fact useful proxies for detecting patterns of anthropologically interesting processes of artifact production, selection, and movement. Combinations of these proxy patterns from archaeological assemblages can further inform the broader behavioral configuration of hominin groups over the landscape that led to the formation of the archaeological record.

1.5 The Structure of This Dissertation

Ultimately, the question is how we, as archaeologists, arrive at the end goal of archaeological interpretation. If, instead of following a series of small steps constructed upon sound inferential practice, we leap directly to higher levels of interpretation for the sake of addressing more evolutionary interesting questions, then our explanations will quickly become difficult to substantiate and verify. This implication is further

exacerbated by the role of archaeology in the multidisciplinary field of paleoanthropology, where interpretations from various subfields are constantly being referenced to test and construct models of human evolution. Returning to the four domains of archaeological knowledge outlined earlier, it appears that constructing the various behavioral and cognitive inferences for understanding hominin evolution is not as straightforward as it seems. This is not to say that they are not achievable, but rather one must follow a chain of sound inferential logic based upon uniformitarian principles with clear consideration of the nature of archaeological categories and the structure of the archaeological record.

This dissertation, a collection of three published or publishable articles, represents an effort to explore the potential of developing lithic archaeology in to an archaeological science project. The first paper (Chapter 2) is a manuscript that explores the nature of archaeological inference creation and the role of experiment in stone artifact archaeology. Specifically, the chapter critically examines the nature of conventional lithic experiments with respect to experimentation as a scientific method of variable testing. It is argued that the emphasis on replication as a common goal of lithic experiments causes the underlying reasoning to suffer from various inferential problems that are difficult to reconcile in a scientific manner. Instead, a focus on lithic experiments should be shifted towards the control of variables and the assessment of baseline patterns that can be unequivocally attributed to the controlled variables.

The second paper (Chapter 3) is an article that has been previously published in *American Antiquity* (Lin et al. 2013). It serves as an example that uses basic fracture properties of stone to detect patterns that can be related to higher level inferences of technological behavior. The paper examines the use of a highly controlled experimental set-up to evaluate the relationship between flake platform attributes and the distribution of flake edge versus volume. This latter property of relative flake edge is defined as a measure of flake utility. By mapping the interrelationship between platform variables, a model is developed to trace changes in the configuration of flake utility and economization among archaeological assemblages. A test case study is presented that is based on lithic assemblages from three Middle Paleolithic sites in the Dordogne region of southwestern France.

The third paper (Chapter 4) is a manuscript currently under review in the *Journal of Archaeological Science* (Lin et al. submitted). It serves as an example of using controlled flintknapping experiments and statistical procedures to apply a geometric index to lithic assemblages as a measure of artifact transport. The paper employs a cortex quantification approach on lithic assemblages from three Middle Paleolithic sites in southwestern France. Flintknapping experiments and statistical approaches of bootstrapping and Monte Carlo sampling are used to establish statistical significance for the calculated Cortex Ratios. Variations in the cortex proportion among the study assemblages over time are considered with respect to the possible shifts in Neanderthal movement pattern associated with environmental changes during the late Pleistocene.

1.6 Conclusion

Archaeology is one of the fundamental subfields within the discipline of paleoanthropology. The rich material record preserved through time provides archaeologists a wealth of information for understanding the behavioral aspects of hominin evolution and population history. However, the process of relating material remains, in this case stone artifacts, to hominin behavior involves various theoretical challenges that are often left implicit in current archaeological research. These theoretical issues have significant metaphysical implications for the conduct of archaeological research and the integrity of archaeological interpretation. Thus, it is not only necessary but critical for archaeologists to confront these theoretical issues in the 21st century in order to move stone artifact archaeology forward as a scientific field. Here, science does not refer to a strict positivist and deductive position of inquiry. Instead, it means having a clear and explicit understanding of its subject matter and the ability to produce meaningful inferences about the past based on solid inferential grounding (Binford 2001).

Furthermore, as various subfields within paleoanthropology have become increasingly integrated and are marked by greater collaboration and communication, the influence that archaeological knowledge will have on our understanding of human evolution will be greater than ever. Therefore, it is equally, if not more, important for the theoretical issues surrounding the production of archaeological knowledge to be transparent and available to non-archaeology specialists, as this will create the foundation for further evaluation and communication among fields as paleoanthropology continue to grow as a true multidisciplinary discipline.

CHAPTER 2: Experimental Design and Inference Building in Stone Artifact Archaeology[1]

2.1 Abstract

Lithic researchers rely heavily on experimentation to infer about past behaviors and activities based on stone artifacts. Yet, discussions of the background method and theory of experimentation and its relation with archaeological inference building continue to be lacking in current literature. This paper explores the analogical nature of archaeological inference and the relationship between experimental design and inference validity in stone artifact archaeology. Conventional replicative lithic experiments lack vital aspects of scientific experimentation, and thus are plagued by inferential issues of analogical adequacy and confidence. It is argued that a greater emphasis on variable control in experimental set-ups is needed in order to establish sound referential linkages upon which constructive analogic inferences about the past can be built.

2.2 Introduction

Experiments have played a central role in the development of lithic studies. In the late 19th century, experimental replication of prehistoric stone artifacts was used to demonstrate their anthropogenic origin, and, by extension, the antiquity of humankind (Johnson 1978). During the 1950s and 1960s, the work of Crabtree, Bordes, Tixier and others further brought lithic experiment to the forefront of lithic studies. In the following decades, the field saw a surge of experimental studies exploring various aspects of lithic

---

[1] Author(s): Sam C. Lin.

technology through the replication of stone artifact forms, from basic properties of fracture mechanics and flake formation (e.g., Cotterell and Kamminga 1987; Cotterell et al. 1985; Dibble and Whittaker 1981) to lithic variability related to percussion techniques (e.g., Barham 1987; Flenniken 1987; Kobayashi 1975; Newcomer 1975; Sollberger 1985; Speth 1974), reduction strategies (e.g., Amick et al. 1988; Flenniken 1978; Newcomer 1971; Sollberger and Patterson 1976) and resharpening (e.g., Flenniken and Raymond 1986). Research areas also expanded from the production of stone artifacts to topics of function, use, and efficiency (e.g., Crabtree and Davis 1968; Fischer et al. 1984; Hayden 1979a; Kamminga 1980; Sheets and Muto 1972; Walker 1978).

Today, lithic experiments come in a variety of forms involving different research designs, methods, and questions. Increasingly, studies also use experimental reconstructions to address issues beyond the immediate production and use of artifacts, including the identification of technical systems, end-product design and production (e.g., Boëda 1993, 1994, 1995; Boëda et al. 1990; Delagnes and Meignen 2006; Meignen et al. 2009; Mourre et al. 2010; Scimelmitz et al. 2011), skill and knowledge transmission (e.g., Eren et al. 2011; Geribàs et al. 2010; Nonaka et al. 2010), and the potential selective pressure on hominin cognition and biomechanics associated with the habitual production and use of stone tools (e.g., Key and Lycett 2011; Stout et al. 2000, 2014; Williams 2011; Williams et al. 2012).

However, as researchers increasingly rely on experiments to draw higher order inferences of hominin behavior and evolution from stone artifacts, much less attention

has been paid to the nature and design of lithic experiments and their relation to the generation of archaeological inference. Specifically, while the methodology of experimentation is powerful, the confidence and security of the resulting inference is strongly contingent on the design of the experiment as well as the underlying analogic premise. To be sure, this sort of inquiry into the nature of archaeological reasoning and the role of experimentation is not new. For example, the rise of the New Archaeology in the 1960s and 1970s led a number of archaeologists to shift their focus to the philosophy of science (e.g., Fritz and Plog 1970; Watson et al. 1971). This shift was largely driven by the desire to establish firm referential frameworks, consisting of 'middle-range theories', to connect the static archaeological record to the dynamic yet unobservable past (Binford 1962, 1977b, 1981; also see Raab and Goodyear 1984; Schiffer 1998). Among these discussions, the role of experimentation was emphasized as part of the 'hypothetico-deductive' process for testing and falsifying existing assumptions of archaeological interpretation (Ascher 1961b; Schiffer 1975). When a hypothesis resists falsification under experimental testing, it is viewed as potentially valid in the sense that the underlying principle can continue to be used for drawing archaeological inference until falsified by further testing (Outram 2008). In this context, validity can be defined as "the best available approximation to the truth or falsity of propositions" (Cook and Campbell 1979:37).

From this perspective, experiments serve as one of the 'gatekeepers' for determining whether certain sets of knowledge in archaeological interpretation can be substantiated by empirical data. Given this important role, however, it is ironic that there

appears to be a lack of discussion about the basic attributes of experiments and the ways in which they should be designed to test hypotheses (Bartovics 1974). In his original publications of *Archaeology by Experiment* (1973) and *Experimental Archaeology* (1979), Cole devised a series of rules for the design of archaeological experiments to ensure a general level of inferential rigor and reliability. While these rules serve as useful general guidelines, their connection to the broader theoretical framework of archaeological inference generation remains under-developed. For example, Cole (1973:18) stated that "[t]he experiment should be assessed in terms of its reliability, that it asked the right question of the material, that the procedure adopted was appropriately conceived…" However, it remains unclear exactly how experimental reliability and research questions are connected, and how their appropriateness should be assessed in relation to the adequacy and validity of the resulting inference.

Recently, studies have increasingly recognized the importance of experimental design and its implication for the adequateness of the resulting archaeological inferences. In zooarchaeology, for instance, Domínguez-Rodrigo (2008) demonstrated that variation among the outcomes of experiments more likely reflect differences of assumptions in experimental design and the underlying analogic premise rather than the studied material per se. This observation has direct implications for not only the inferential security but also applicability and comparability of experimental interpretations for past human behavior. This issue is especially pressing as archaeological knowledge is progressively being sought by disciplines such as genetics and physical anthropology to validate and contextualize models of human evolution and population dispersal (Henke and Tattersall

2007). It is within this context that the need for researchers to critically consider the theoretical interaction between experimental design, research strategy, and archaeological reasoning becomes all of the more imperative.

Extending such observations to lithic studies, the goal of this paper is to examine the properties of experimental design and their interaction with the creation of archaeological inference of stone artifact archaeology. In order to do this, however, it is first necessary to consider the nature of archaeological reasoning in the form of analogic argument, particularly in terms of 'formal' versus 'relational' analogy. We then compare flintknapping as an actualistic approach to scientific experimentation with respect to the issue of confounded variables and the ways that uniformitarian assumptions are treated in the formulation of hypotheses. This analysis is accompanied by discussions regarding the basic properties of experimentation, including variable control, sources of error, and inference validity. Finally, we differentiate between 'pilot' versus 'second generation' experiments and discuss their respective roles in the process of archaeological knowledge generation.

2.3 Analogy in Archaeological Reasoning

According to Gibbon (1989:142-172), the structure of social science inquiry can be separated into three main realms: 'the observed', which is the phenomena we perceive from the world; 'the empirical', which is the data we construct from the observed; and 'the real', which is the actual condition or process we wish to understand through the analysis of the empirical. In many disciplines, the interaction between the observed and

the real operates at a temporal scale that is discernible within an individual's lifetime. For these fields, general theories regarding the phenomenon in question can be derived from multiple observations and verified through experimentally replicating the relevant processes and variables. For archaeology, however, the processes that led to the formation of the material outcome which we observe today (the archaeological record) cannot be directly experienced. Unless there are other sources of information, such as historical documentations, we rely instead on analogy to make inferences about the past by linking concepts and relationships derived from the present to aspects of the archaeological record (Binford 1981; Gifford-Gonzalez 1989; Wylie 1982, 1985).

In her seminal paper, Gifford-Gonzalez (1991; also see Wylie 1985) differentiated between the use of 'formal analogy' and 'relational analogy' in archaeology. Based largely on the ordinary experience of the observer, formal analogy operates by drawing a causal connection between observed modern process and its material outcome. When archaeological items share formal similarities with a modern object, it follows that the observed process in the contemporary world also occurred in the past. The operation of formal analogy is summarized schematically by Gifford-Gonzalez (Figure 1) and contains three key assumptions: 1) the linkage between the observed modern object and process is causal; 2) the similarities between the modern and prehistoric objects are meaningful to the analogical inquiry; and 3) the inferred process is uniformitarian in nature.
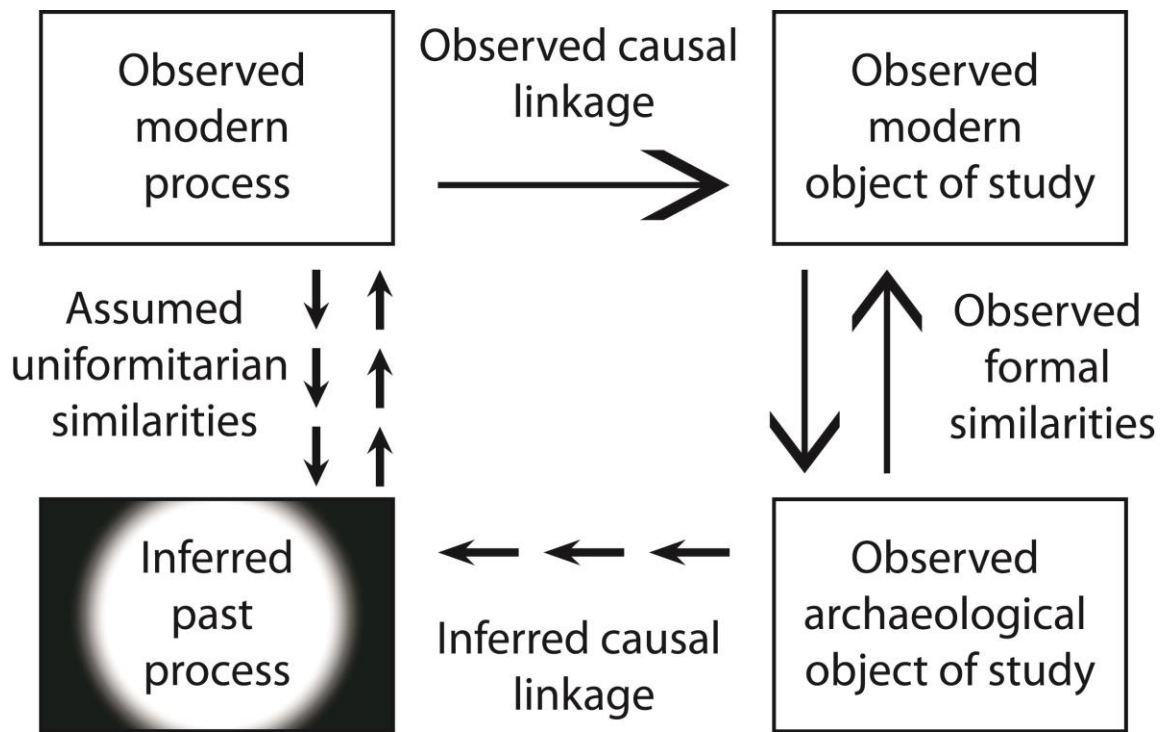
Figure 1 – Gifford-Gonzalez's model of analogical reasoning in formal analogy. Blank boxes indicate observations made in the present world. Redrawn after Gifford-Gonzalez (1991:222).

Most of archaeological systematics is rooted in formal analogy. Because the archaeological record is an anthropological phenomenon, it is intuitive for archaeologists to use their daily experience of the human world as primary analogs for making sense of archaeological remains. This form of analogy can be seen in the basic naming and description of artifact categories to interpretations of past events and processes. The most explicit demonstration of formal analogy is the early approach of ethnographic analogy, where living groups are chosen as counterparts of past societies. The selection conditions for ethnographic analogs involve formal similarities in either material culture or ecological and subsistence conditions (Ascher 1961a; Stiles 1977).

Two inferential issues concerning formal analogy have been identified. The first is that an inference is primarily a projection of modern knowledge into the past, and thus what archaeologists can infer is inherently restricted to what is already known from the present (Gould and Watson 1982; Gould 1980). As such, the reasoning itself is incapable of generating general theories or novel knowledge about the past. It is, therefore, unclear how archaeologists can identify phenomena that are unknown in the present-day world, or whether or not it is even possible. Second, formal analogy operates more as a logical statement, and offers no obvious means to gauge the security of the derived analogic inference (Gould 1980). In some ways, an analogy is considered valid by default as long as the analog shares with the study subject similarities that are deemed relevant. As a result, considerations of inferential security come to rest on the exhaustive listing of selection criteria for modern analogs, or philosophical debates over the assumptions of causality with respect to the phenomenon under study (Ascher 1961a; Binford 1967; Stiles 1977). This inability to ascribe inferential confidence led to the dilemma of determining how far archaeological interpretations could and/or should be taken at the cost of methodological rigor (Wylie 1985).

Then again, the trade-off between inference confidence and methodological rigor only becomes apparent as one moves from empirical 'facts' of archaeological materials to contextual interpretations of past practices. This concept of a hierarchical order of inference is exemplified by the work of Hawkes (1954), who postulated that, as one moves up the inferential order away from phenomenon restrained by the natural, physical world (i.e., technology, subsistence, and economy) to those of the socio-political and

ideological realm (i.e., intentionality, social norm, cultural tradition), the reliability of the inference drops significantly due to their greater reliance on culturally specific manifestations. From this perspective, the issue of inferential security in formal analogy can, in part, be attributed to the lack of differentiation between different inferential orders (Wylie 1985). Indeed, under formal analogy, all inferences are constructed upon a singular, causal relationship, regardless of the differences in the number of contextual variables involved among them. Consequently, the relationship between these variables and the study phenomenon are confounded and become impossible to tease apart.

In contrast with formal analogy, relational analogy differentiates the inferential process into individual linkages that are structurally organized on the basis of a body of referential knowledge. This body of referential framework is internally coherent and therefore allows the construction of analogic inference by connecting multiple justifiable inferential linkages. Binford (1962, 1977b, 1981) conceptualized these referential linkages as 'middle-range theories'. All middle-range theories need to have the following attributes (Wandsnider 2004). First, the causal connection between the material phenomenon and the generating process has to be *unambiguously* demonstrated and documented. Second, this causal connection has to be of uniformitarian nature and, more importantly, can be warranted as such, so the inference can be projected into the past with warrantable confidence. To be clear, these two attributes already exist in the operation of formal analogy (as illustrated above in Figure 1), although they tend to exist in implicit and confounded forms embedded within the underlying assumptions. The third attribute is that middle-range theory has to operate independently of ideas about the past that

archaeologists wish to investigate, and thus can serve as a neutral medium for inferring the occurrence of past processes from the archaeological record.

While the definition and utility of middle-range theory has been debated and critiqued since its original proposal, many scholars have arrived at the same conclusion that in order for any archaeological inference to hold some level of security, the involvement of referential knowledge in forms similar to middle-range theory is critical (Trigger 1995; Tschauner 1996). Binford (1981, 2002; also see Wandsnider 2004) later termed the systematic compilation of these established referential knowledge as "frames of reference". To be sure, the point of discussion here is not to equate relational analogy with the Binfordian notion of middle range theory and frames of inference. Instead, the goal is to characterize relational analogy as a separate analogical reasoning process with specific properties that are different from formal analogy. Building from firm referential linkages, relational analogy is considered to be more strongly warranted than formal analogy (Binford 1981; Gifford-Gonzalez 1991; Wylie 1985). Its explicit treatment of uniformitarian principles also forces researchers to confront the assumptions that underlie inference construction.

2.4 Experimentation in Lithic Studies

The goal of establishing firm referential linkages for connecting aspects of the archaeological record to past dynamic processes, led many researchers in the 1970s and 1980s to conduct experimental studies on various topics of archaeology. However, contrasting with the conventional definition of experiment as a scientific method, these

archaeological studies align more with actualistic research of discovering and evaluating archaeological relevant variables within living contexts. In fact, in one of the original articulations of middle-range theory, Binford (1981) advocated for archaeologists to conduct actualistic studies for detecting possible referential knowledge. His extensive ethnoarchaeological work with the Nunamuits constitutes a primary example of this sort of effort (Binford 1977a, 1978a,b, 1979, 1980, 1981). The source of this emphasis on actualistic research can be attributed to the analogical nature of archaeological reasoning. Because past processes can only be comprehended through modern observations, it makes sense that a general understanding of these processes can be gained through replicating past activities and behaviors.

This sentiment is particularly apparent in stone artifact archaeology. Since few societies today use stone tools on a daily basis, lithic technology and stone artifacts in general remain largely a body of alien knowledge to modern archaeologists. While ethnographic accounts provide valuable information regarding the ways stone tools were used in modern living contexts, their direct applicability to archaeological material remains limited as they often represent a subset of technological behavior operating within a specific cultural context and time scale. As such, stone artifact researchers have come to rely on experimental replication as a principle means to understand the behavioral processes that underlie the formation of stone artifacts in the archaeological record. They have done this by primarily replicating particular artifact forms and their associated production procedures through flintknapping. The concept of replication in this context has been defined by Flenniken (1984) as the consistent recreation, with the

same lithic materials, the same reduction technology and end products as the prehistoric knapper. This anthropological sense of 'replication' is often distinguished from 'flintknapping', which denotes only the effective production of flaked stone artifacts. The emphasis on systematic and technical replication also places demands on the relative skill level and expertise of the knapper who performs the replication experiment, as well as his/her familiarity and knowledge of the artifact that is being replicated (Flenniken 1984). The results of replication are then compared with archaeological materials on specific criteria to determine the analogic validity of the experimental technique for inferring reduction practices in the past (Mourre et al. 2010).

The goals of experimental replication have shifted considerably since the initial adoption of the approach in lithic studies. Prior to the 20th century, scholars used replicative flintknapping primarily to support the artifactual nature of ancient chipped stone implements. This was done in part due to the recognition that flakes can be produced under natural conditions and thus the need to discriminate cultural artifacts from naturally flaked stones (Evens 1872; Skertchly 1879; also see Johnson 1978; Lerner 2013;). In the late 19th century, Holmes (1894) employed replicative flintknapping in his study of American bifaces as a way to gain an understanding of the production sequence of artifacts from raw material acquisition, technical production, to forms recognized archaeologically. This notion of sequential manufacture to arrive at a particular end-product became a key element of modern lithic experimentation (Bleed 2001).

In the first half of the 20th century, increased attention was given to experienced flintknappers and the documentation of various knapping techniques for replicating artifact types that are comparable to those observed archaeologically (Johnson 1978; Lerner 2013). Flintknapping replication was further popularized and incorporated into the realm of mainstream archaeological investigation in the 1960s by the work of several skilled flintknappers, including Crabtree, Bordes, Tixier, Callahan, and Bradley. Increased consideration was also paid to the reporting and/or control of variables involved in knapping experiments, such as the size and nature of the raw material, the technique of production, and the properties of hammerstones used (Johnson 1978). However, much of the literature on lithic replication during this time focused on identifying the how-to or craft aspects of replicative flintknapping rather than answering specific archaeological questions (Andrefsky 2005).

Since the 1980s, the focus of experiments shifted from end-product manufacture to the production of by-products as well as the overall reduction sequence. Specifically, studies examined the interrelationship between different knapping sequences with the characteristics of the overall produced lithic assemblage. This is best represented by the reduction sequence approach of North American archaeologists (Bleed 1996; Dibble 1984, 1987, 1995b; Frison and Raymond 1980; Morrow 1997) and the *chaîne opératorie* school of France and Continental Europe (e.g., Boëda 1986, 1988; Boëda et al. 1990; Geneste 1985; Perlés 1992; Pigeot 1990; Sellet 1993). These studies stress the importance of the technological/behavior process that underlies the formation of lithic artifacts, which could effectively be captured through refitting and replicative

flintknapping experiments. One of the main objectives of this approach to replicative experiment is to transcend traditional typological units and holistically investigate the procedural steps in the entire manufacture sequence of finished products (Bleed 2001). Experiments within this context are done by consistently replicating not only specific artifact forms but the entire manufacturing sequence and the technical features of the associated knapped products. The presence of reconstructed reduction strategies can, in turn, be identified in archaeological assemblages through either refitting or diagnostic artifacts that carry markers of specific reduction sequence (e.g., Boëda et al. 2013; Li et al. 2009).

Regardless of the different goals of experimental replication, the inferential basis of the approach differs markedly from that of conventional scientific experimentation, and instead aligns more closely with formal analogy described earlier. Namely, a single, causal relationship is stipulated between a specific material outcome (e.g., a particular type of flake or flake attribute) and a generating process (e.g., particular reduction technique) based on the ordinary experience and perception of the observer regarding the phenomenon of interest. While the inference may appear to be supported by the matching qualities between modern experimental materials and prehistoric artifacts, its validity is, in fact, unverifiable due to the nature of the analogy. If an analog and the phenomenon in question share qualities that are considered applicable to the research inquiry, then the analogy is taken as valid; if the two subjects do not share relevant similarities, then the analogy is naturally rejected.

To further illustrate the dissimilarity between lithic replication and experimentation as a scientific method, it is useful to compare the nature of the two approaches. According to Bartovics (1974:201; also see Kosso 2011), experiments should contain four essential elements:

1) They require a hypothesis that stipulates a connection between the observed material outcome and the generating conditions and processes for the phenomenon.

2) Their set-up needs to allow a specific degree of manipulative control over to variables involved in the phenomenon under study.

3) They must rely on objective, and preferably quantitative, modes of documentation in order that the experimental outcomes can be reproduced by other investigators.

4) They are empirical demonstrations rather than arguments or debates.

While all four attributes are critical for a strong experimental setup, the validity of an experiment is primarily determined by the first two elements.

2.41 Hypothesis Construction and the Treatment of Uniformitarian Assumptions

To carry out an experiment, the researcher must first have some idea of how the study subject is connected to the phenomenon in question. The sources of these ideas may vary. They could be derived inductively through particular observations, or

abductively through the logic of best probable explanation (Niiniluoto 1999). For these ideas to be turned into hypotheses, they have to be stated in an explicit cause-and-effect statement that is falsifiable. If stated correctly, an opposing hypothesis (i.e., null hypothesis) that postulates the neutral or null expectation of the original hypothesized relationship would naturally exist. As discussed above, the goal of an experiment is to test whether the null hypothesis can be successfully rejected. If so, then the stated hypothesis is considered to be valid in the sense that the postulated relationship can continue to be used for inference construction until it is falsified by further testing.

The key to a good hypothesis is its capacity to be rejected. This characteristic may seem straightforward, although, it is often more difficult to achieve due to the ways the underlying assumptions and premises are outlined. Specifically, all experiments require uniformitarian assumptions in order to generalize particular observations to other settings. However, whether a hypothesis is falsifiable is, in part, dictated by how the underlying uniformitarian theory is treated. According to Bailey (1983), when a uniformitarian theory is treated 'methodologically', it becomes a means for assessing *another* theory. When a uniformitarian theory is treated 'substantively', on the other hand, it becomes a substantive extension of the theory that is to be investigated (c.f. Gould 1965). It can be said that all archaeological inquiries are driven at some level by substantive uniformitarianism. From biological to socio-cultural evolution, cultural transmission, and behavioral ecology, these theoretical frameworks that archaeologists employ require some level of uniformitarian assertion to warrant investigation of related topics in the past.

However, if the uniformitarian theory underlying the test hypotheses is substantively treated, the test conclusion is bound to be predetermined by the preexisting theory (Bailey 1983). For example, if a hypothesis asks whether certain flake forms were manufactured as end-products by a specific reduction strategy, the test outcome is necessarily predetermined by the underlying uniformitarian assertion – that certain flake forms represent end-products and that the knapping process, irrespective of time and space, is driven by the production of these end forms. What is problematic here is that the question that archaeologists wish to investigate is already presumed, implicitly, by the uniformitarian theory. Thus, if the knapper successfully produces the specific flake forms with the reduction strategy, the experimental technique is taken to be valid for inferring reduction practices in the past. If a knapper fails to produce these flake types, on the other hand, then it is either due to the use of incorrect knapping strategy or that the knapper's skill is insufficient. One way or the other, the hypothesis of whether the specific flake forms can be manufactured as end-product becomes an untestable assertion, given that the uniformitarian nature of these flake types as end-products are already asserted substantively.

Another danger in the substantive treatment of uniformitarianism is the risk of generating tautological arguments. Since these kinds of hypotheses cannot be falsified, the study outcome could only be taken to substantiate the original assumed premise. As a result, what archaeologists can learn about the past is inherently limited by the analytical categories which we have generated. If studies investigating past knapping behavior begin by asserting that the ways with which knappers reduce stones are uniformitarian

44

and, hence, can be modeled as such on the basis of ethnographic or actualistic analogs, then the resulting inference is bound to be expressed in these predefined ways. In other words, the limiting factor in archaeological inference is not the archaeological data themselves, but rather the boundaries of our own imaginations about how stones could be reduced (discussed further below).

However, if we treat the uniformitarian theory methodologically, the process with which we assume to operate in the present as well as in the past is required to be intellectually *independent* from what we wish to investigate (Bailey 1983; also see Binford 1981). This treatment grants independence to the formulated hypothesis from its theoretical origin, and constitutes a chief merit of scientific reasoning (Bettinger 1987; Hempel 1965). If, for example, we are interested in determining whether certain flake forms are made as end-products, we may first ask whether certain reduction strategies can produce flakes that contain the relevant morphological characteristics as the archaeological artifacts of question. This inquiry is based on the uniformitarian assumption that fracture mechanics and stone property operate in an invariant way, and thus mechanical processes that led to the creation of specific flake types observed in the present very likely also operated in the same way in the past. At this point, researchers may examine whether these specific flake types were treated in any ways different from other flake forms, such as in their use, resharpening, and transport (e.g., Dibble 1995a). These results in turn help support whether the artifact forms were indeed end-products that were preferentially selected and manufactured, or, on the other hand, if they share no clear distinction with the rest of the assemblage and perhaps represent arbitrary forms

singled out by modern archaeologists (Bar-Yosef and Dibble 1995; Bar-Yosef and Van Peer 2009; Dibble 1995a).

It is useful to point out here the importance of keeping independent the criteria for assessing differences between artifact groups from the attributes used to define these categories in the first place. Otherwise, the resulting pattern would risk reflecting the definitions of these categories made by the researcher more than the actual reality and difference between the groups. For example, Schlanger (1992) and Eren and Lycett (2012) examined the difference between Levallois end-products with non Levallois debitage flakes based on formal attributes (Schlanger used dimensional measures; Eren and Lycett used geometric-morphometric analysis). Because the differentiation between the two artifact groups is defined in part on the basis of formal characteristics (Levallois flakes as defined are generally larger, wider, and relatively thinner), it was unclear whether the observed difference between the two categories was meaningful. A similar point regarding the reality of ovate versus pointed handaxe was discussed by McPherron (1999). The point is that the uniformitarian theories for establishing our inferential reasoning and analytical units for conducting these studies should be independent from the research question asked, and, more importantly, that the hypothesis can only be falsified or supported, but cannot be proven to be true.

A methodological uniformitarian approach for formulating hypotheses requires that the phenomenon under experimental testing be unrelated to human behavior. This may seem counterintuitive. After all, the business of archaeologists is to explain the

archaeological record in behavioral terms. If the test hypothesis is required to explain

processes that are independent from our research goal, then the inference resulting from

the experiment would also be independent from behavior. For example, the connection

between exterior platform angle, platform depth, and flake size (Dibble and Rezek 2009;

Dibble 1997; Lin et al. 2013) is, in and of itself, independent from anthropogenic

activities – i.e., the processes operates regardless of hominin behavior. It is simply a

physical phenomenon that operates in the natural world. However, such relationship,

verified through experimentation, can be effectively applied to archaeological material to

detect patterns from which further behavior interpretations can be drawn. This is an

important departure from the Binfordian approach of middle range theory, where the

referential linkages for constructing relational analogy are sought in human terms and

generally situated in a behavioral ecology framework (Gifford-Gonzalez 1991). Such

emphasis on the theoretical reduction of uniformitarian behavior has instead yielded

trivial or oversimplified explanations (Flannery 1973). If the research question is instead

framed on a methodological uniformitarian basis, then the interpretation of past behavior

can be constructed on patterns that are neutral and, more importantly, verifiable through

experimentation.

2.42 Experimental Control and the Issue of Confounding Variables

In an experiment, a hypothesis is operationalized into variables that are relevant to

the inquiry. In general, an experiment manipulates one or more <u>independent</u> variables and

then records changes in the <u>dependent</u> variables while exerting control on *all* other

<u>nuisance</u> variables (Kirk 2009, 2012). At this point, it is worth emphasizing that it is

pivotal for the various independent variables involved in the phenomenon under study to be explicitly parsed out. This is done to avoid possible confounding relationships between independent variables where the levels of one independent variable are systematically associated with the levels of the other (Abdi et al. 2009:6). When two or more independent variables are confounded it, becomes impossible to interpret the results because the outcome may be caused by one variable or the others, or from the interaction among them (Figure 2). For example, for a study concerned with whether flake edge angles affect cutting efficiency on wood, independent variables would include edge angle, edge length, flake thickness, flake weight, wood quality and consistency (e.g., hardness), cutting force, cutting speed, and cutting time. If, for instance, flake edge angle and flake size are not separately treated and the two variables are systematically associated (e.g., larger and heavier flakes have on average higher edge angle), then the effects of the two independent variables will be confounded, making it impossible to know if the outcome reflects the influence of edge angle, or flake weight, or the combined effect of the two variables.
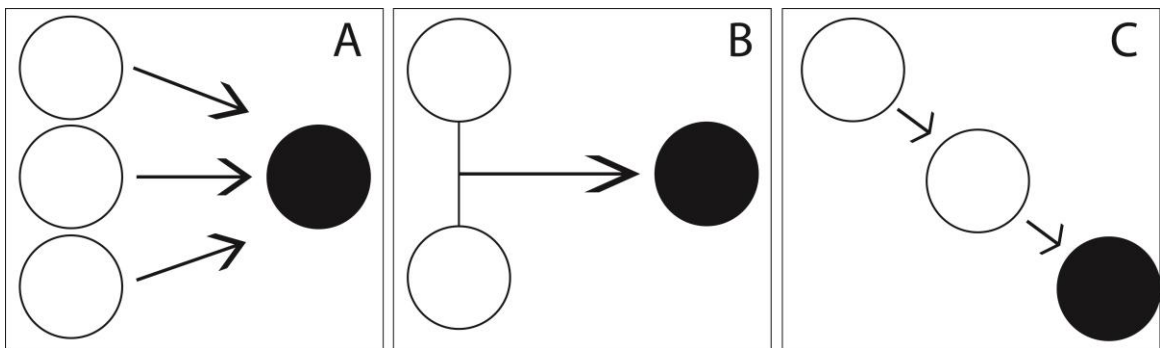
Figure 2 – Scenarios of confounding variables in experiments. White circles represent independent variables; black circles represent dependent variables. Scenario A: the dependent variable is simultaneously influenced by multiple independent variables. Scenario B: the dependent variable results from the

interaction between two (or more) independent variables. Scenario C: the causal linkage between the dependent and independent variables is triggered by other independent variable(s).

Actualistic studies, including replicative flintknapping, are particularly prone to the problem of confounded variables. As many flintknappers would agree, flintknapping is a complex process involving a wide range of variables – from fracture mechanics to nodule geometry, individual technique, skill, intention, technical know-how and folk knapping wisdom, as well as knappers' biometric capacity and cognitive ability to decide, plan, and execute specific knapping actions. In replicative flintknapping, many of these independent factors are allowed to vary freely during the knapping process. Idiosyncratic characteristics due to the use of human participants also potentially further introduce confounded variables that systematically bias certain variables (Kirk 2009).

As an example, Whittaker (1994) observed unique regional patterns in North American knapping circles where knappers employ distinct techniques to achieve similar end goals. Knappers, thus, are likely to systemically associate certain historically-derived practices, almost at a subconscious level, with particular knapping activities. As such, it becomes impossible to attribute experimental outcome directly to specific variables. To be sure, most replicative studies do exert some form of control over specific test variables. Thus, studies often hold some level of control over raw material, hammer type, and knapping techniques throughout the flintknapping process (e.g., Amick et al. 1988; Henry et al. 1976; Newcomer 1971). Then again, because the extraneous nuisance variables involved in the knapping process remain uncontrolled, their relationship with the study outcome remains confounded. It is also worth noting here that monitoring the changes of these uncontrolled variables also does not resolve the confounded nature of

49

the experimental outcome. Without actual control, it remains impossible to discern whether the relationship between these monitored parameters and the test outcome is indeed causal and meaningful.

In light of these issues, a number of studies turned to the use of artificial setups with flaking apparatus to simulate the process of flake formation under truly controlled conditions (Bonnichsen 1977; Dibble and Pelcin 1995; Dibble and Rezek 2009; Dibble and Whittaker 1981; Dibble 1997; Faulkner 1972; Pelcin 1997a,b; Speth 1972). The emphasis in these studies tends to be placed on the ability to vary relevant variables while holding others constant throughout the experimental process. This is done with the goal of explicitly testing the interactions between specific independent variables under the control of the flintknapper and other variables that are observable on flakes (Dibble and Whittaker 1981). For example, recent studies by Dibble and colleagues (Dibble and Rezek 2009; Lin et al. 2013; Magnani et al. 2014; Rezek et al. 2011) employed highly controlled settings and examined the relative effects of core surface and platform morphology and the various modes of force application on flake attributes using molded glass cores. This sort of study, termed a "controlled experiment" by Dibble and Whittaker (1981), is designed specifically to manipulate and control variables involved in specific aspects of flake formation for the goal of discerning their respective effects on flake attributes. This approach contrasts with 'replicative experiment', which emphasizes the replication of artifacts by flintknapping under settings that resemble past knapping conditions.

However, experiments performed under more controlled settings are not necessarily free of confounded variables. Indeed, early controlled studies by Speth (1972) and Bonnichsen (1977) produced confounded results due to the lack of separation between systematically associated variables, such as hammer mass versus velocity, and the angle of hammer blow versus exterior platform angle (see Rezek et al. forthcoming). Then again, under controlled settings, the chance for confounded variables to occur is less as the control of the various variables is more explicitly delineated. Confounded variables are not always easy to identify and may require repeated experiments to untangle the interaction between correlating variables. One effective approach for identifying confounded variables is through duplicating the experimental setup (Abdi et al. 2009). Given that all variables are treated in the same way, if the replicated study does not produce the same results as the original, then it is reasonable to suspect the presence of confounded variables.

Studies are also more prone to confounded variables when they employ nominal classifications of independent variables, which are groupings representing predefined conditions and therefore cannot be manipulated by the experimenter (Abdi et al. 2009). Examples of these in lithic studies include core versus flake, flake versus tool, or end-product versus debitage. Objects assigned to these classificatory groups would inevitably be different in many other respects and can therefore confound the possible interactions between independent variables. For instance, soft and hard hammer are predefined categories that an experimenter can employ as a variable distinction. However, hammers belonging to either one of these classifications are inevitably different in many other

ways (morphology, weight, hardness, elasticity, grain size, grain morphology, etc).

Without further separation, the differences in knapping results between the two hammer

material types can only be attributed to the level of the nominal classification (Magnani et

al. 2014).

2.43 Inferential Validity, Sources of Error, and Equifinality

At the end of the day, the question of experimental design and inferential validity

boils down to whether or not the experiment setup is sufficient in answering a specific

research question. When only a single independent variable is considered, the

experimental design is often straight forward. However, such limited experiments also

tend to be further away from any natural setting due to their simplistic design.

Experiments can become more realistic when the number of independent variables is

increased to approach reality. In the social sciences, this is referred to as an increase in

'external validity,' or where the experimental results are valid beyond the limits of the

experimental setting and can be generalized and applied across a wide array of contexts

(Abdi et al. 2009; Campbell and Stanley 1963; Kirk 2009, 2012). The contrast to this is

'internal validity' which depends on the precision of the experiment within the

experimental context, where it is possible to conclude accurately that an independent

variable is responsible for variation in the dependent variable of question.

By this definition, internal validity is inversely correlated with the size of

experimental error, which represents the total variability in the dependent variable due to

causes other than the tested independent variables (Abdi et al. 2009:11). Internal and

external validity thus is commonly seen as the two opposing ends of a continuum that stresses simplicity and clarity on the one hand and complexity and generality on the other. At one level, it is useful to think of controlled experimentation as high in internal validity for its internal consistency within the experimental setup due to its high level of control over all variables. Flintknapping is high in external validity due to the number of variables involved in the process and its proximity to the reality of the knapping process. However, given that the reality of past processes cannot be intuitively attained, viewing external validity as an approximation of past reality is problematic. Alternatively, a more productive position may be to view external validity as the degree of generalizability of the observed inference to other settings. From this perspective, external validity decreases if the experiments focus more narrowly on the particularities of past lithic technology, regardless of whether the experiments are performed in a controlled fashion (e.g. production of Mesoamerican prismatic blades by Faulkner 1972) or by flintknapping (e.g. Levallois flake as the object of study in Eren and Lycett 2012). In this sense, controlled experiments focusing on fundamental fracture principles are also high in external validity because the investigated relationship is relevant to the study of the production of stone artifacts across wide geographical, temporal and technological contexts (Magnani et al. 2014).

Note that the use of internal and external validity here differs from how these terms are used by Lycett and Eren (2013), who argued that the archaeological record is high in external validity due to its empirical reality but low in internal validity because the record is biased and incomplete, due to preservation and sampling (see Sullivan

1978). On the other hand, mathematic models of reduction techniques are high in internal validity for their internal consistency but low in external validity for their abstractness (Brantingham and Kuhn 2001; Brantingham 2010). Experimentation is instead argued as a means for bridging the two data sources to strengthen archaeological inquiry.

However, with this definition, it is difficult to see how the archaeological record or models of flake production can possess any form of validity at all. Specifically, the concept of internal and external validity was developed in the social sciences for the purpose of describing the reliability and applicability of inferences derived from experimentation (Abdi et al. 2009; Campbell and Stanley 1963). Given that an inference is defined as a logical conclusion regarding a phenomenon derived from specific premises (Sullivan 1978), the archaeological record only represents the phenomenon observed by modern archaeologists and does not in and of itself offer any inferential power about past behavior. Likewise, conceptual models of flake production are premises assumed to hold certain relevance to knapping processes in reality. Inference can only be developed when the models are connected with empirical observations and, more importantly, archaeological data. Therefore, Lycett and Eren's (2013) argument seems to reflect a general misunderstanding of the original definition and usage of internal and external validity in the social sciences.

Following the discussion above, a good experimental design that can lead to a secure basis of inferential knowledge in stone artifact archaeology would avoid confounding variables while at the same time possess high internal validity. In other

words, it allows us to determine how confident we can conclude that an independent variable is, in fact, responsible for the variation observed in the dependent variable. In the social sciences, sources of error are recognized that can compromise the internal validity of an experiment. These are 'history', 'maturation', and 'selection'. As Kirk (2009:25) outlines, history describes "events other than the treatment that occur between the time the treatment is presented and the time that the dependent variable is measured"; maturation represents instances where "[t]he dependent variable may reflect processes unrelated to the treatment that occur simply as a function of the passage of time"; and selection is the possibility "that the participants in the experiment are different from those in the hypothesized comparison sample."

For stone artifact archaeology, maturation does not apply as stone artifacts do not change by themselves through time, although this also depends largely on how artifact categories are construed depending on the dominant lithic paradigm. On the other hand, 'history' and 'selection' are potentially problematic for any replicative experiment. In terms of 'history', lithic assemblages represent accumulation of lithic artifacts over time at given locations (produced by many people). The potential complex formation history of lithic assemblages means artifacts found in close spatial proximity may, in fact, share distinct and independent settings of production and use (Dunnell 1992; Turq et al. 2013). In the context of experimental replication, however, these assortments of stone artifacts which archaeologists attempt to replicate through flintknapping are often treated as a collection of materials that are meaningfully related in terms of production sequence. This issue seems difficult to reconcile, given that there exists little empirical evidence

that can substantiate the temporal integrity of lithic artifacts at a production event level. Even refitted artifacts could arguably be produced by different individuals at different points in time.

As a result, the solution to this issue may instead fall on the research question instead of experimental design. That is, if one can make shift away from treating assemblages as systemic collections of products that are connected in some behaviorally meaningful way, then one can develop a view that emphasizes the temporal and processual complexity of assemblage formation (Holdaway and Wandsnider 2008 and references therein; Kuhn 2004; Stern 1994). Of course, this question touches on much broader metaphysical issues concerning the ontology of archaeological thinking, and thus is beyond the scope of the current discussion. The point here instead is to draw awareness to the connection between the conceptualization of archaeological material and experimental design, and its implication for the resulting inference validity.

The second problem, that of 'selection,' relates to the issue of substantive uniformitarianism as discussed earlier: how do modern knappers know if the employed experiment techniques are indeed counterparts to behaviors in the past? In lithic studies, because the knapping process is complex, there exists a common sentiment that the only way to comprehend the process is to immerse oneself in the practice of knapping under conditions, to the best of one's knowledge, that resemble past knapping settings. This belief has often led to an emphasis on ecological validity (Brunswik 1956; originally articulated as "representative design"), i.e., the resemblance to empirical reality in

replicative flintknapping. However, as a historical discipline, archaeology's emphasis on ecological validity appears oxymoronic, since the 'reality' in which past knapping events occurred simply cannot be known. In fact, the accurate reconstruction of past knapping techniques is typically the goal of such experiments. While some level of ecological validity can be assured by using raw materials and hammer materials that are in line with what is found in an archaeological context, various other aspects of knapping conditions are speculative. These include the posture of the knapper (standing, sitting, kneeling), the placement of the nodule (on the ground, freehand, on the lap), the way hammers struck the nodule (thrown, direct percussion, swing-arm percussion). Individual knappers oftentimes preferentially adopt specific combinations of these factors for particular knapping tasks. Though some of these factors may leave discernable traces on flakes, many of these variables remain impossible to detect archaeologically.

Ultimately, the issues of selection and history reflect the question of equifinality in replicative experiment, where distinct configurations of knapping process and formational history can lead to similar material outcomes in artifact morphology and assemblage characteristics (Magnani et al. 2014). Indeed, the process of knapping and stone reduction is sequential and thus can be reconstructed as such in an empirical manner. However, it is more challenging to substantiate that the behavioral processes underlying the formation of the lithic record can be intuitively replicated as systemic knapping sequences.

2.5 Discussion

The inferential issues related to lithic actualistic experiments outlined above is due largely to the replicative nature of the experimental design. Regardless of the theoretical justifications for the criteria of replication, the resulting inference from these experiments is inescapably of formal analogic nature. As a consequence, the validity of the stipulated inference is difficult to verify and in turn has to be taken on faith. Fundamentally, the question of whether our modern replications do, indeed, serve as meaningful analogs to past processes remains unresolvable. Within this context, it is useful to distinguish replicative experiments from others that are orientated to the control of variables in an attempt to strengthen the internal validity of the study outcome (Dibble and Whittaker 1981; Dibble 1997). Unlike replicative experiments, controlled experiments are carried out with the goal of establishing referential linkages upon which relational analogic inferences can be built.

The concept of a controlled experiment is associated almost exclusively with studies that employ artificial settings for flake production (e.g., Dibble and Rezek 2009; Magnani et al. 2014; Rezek et al. 2011). However, this does not necessarily mean that flintknapping as an experimental approach cannot be carried out in a controlled fashion. Although with less ability to control variables comparing to highly controlled setups, flintknapping allows for relatively greater flexibility in variable manipulation and thus provides a useful means for investigating broader topics of lithic assemblage composition and other assemblage-scale variability. However, as previously pointed out, conducting flintknapping experiments in a controlled manner requires much more than just the

control or monitor of few independent variables during the knapping process. Rather, it necessitates a fundamental shift in experimental design from being centered on artifact replication to the controlled examination of the relative effects of specific independent variables. Specifically, the goal of the experiment should be capturing the range of variability attributable to certain knapping factors that operate on uniformitarian processes. By comparing the experimentally derived pattern to the archaeological data, it is then possible to infer the relative effect of the examined variable in shaping the observed archaeological pattern.

The flintknapping process, however, requires a wider range of assertions and inconsistencies over variables that are difficult to control. A possible resolution to this issue is to deliberately maximize the degree of variability in the nuisance variables while maintaining high control over the specific independent variables that are under testing. This approach is employed to minimize the involvement of assumptions regarding how knapping activities took place in the past (the replicative approach). It also increases the internal validity of the resulting conclusion by establishing a sound, cause-and-effect linkage between the tested variable and the experimental outcome. For example, if one wishes to examine the effect of nodule size on assemblage composition, it is more effective to conduct the necessary experiments by employing a range of knapping configurations (reduction technique, hammer type, core placement, and the knapper) while controlling stone nodules at different size levels. More importantly, the pairing of the various knapping configurations to the different nodule size levels need to be arbitrary and random. Such an approach is akin to randomization in experimental design

of the social sciences, and serves to minimize the potential systemic bias of the uncontrolled nuisance variables on the test outcome by randomly distributing their effects across all test groups.

This sort of controlled and comparative design for flintknapping experiment has been adopted by several studies reported in the literature, particularly with respect to artifact reduction intensity. A number of studies in the 1980s and 1990s employed flintknapping experiment to establish models of lithic reduction (Amick et al. 1988; Bradbury and Carr 1999; Ingbar et al. 1989; Magne and Pokotylo 1981; Shott 1996b). These studies controlled reduction intensity at different intervals to examine its relative effect on artifact attributes. Using a similar protocol, Braun (2006; also see Braun et al. 2008) examined the effect of reduction intensity on core attributes by experimentally knapping stones over various iterations of core reduction. The knapping was done independently by multiple knappers having different skill levels. Instead of attempting to replicate the archaeological artifacts, cores were reduced with no specific assumed intention of producing particular core form or end product. Rather, reduction was done with the goal of producing large flakes (Sahnouni et al. 1997). Through multivariate regression analysis, Braun used the experimentally derived pattern of core reduction to draw inferences regarding the relative reduction intensity represented in the Oldowan assemblages in Kanjera South and Koobi Fora. Similar experimental approach to assess the effect of reduction intensity was adopted by Archer and Braun (2010), who examined Acheulian bifaces from the site of Elandsfontein, South Africa, Douglass (2010) for core reduction intensity in the Holocene lithic assemblages from western New South Wales,

Australia, and Marwick (2008) for the flaked stone artefacts from mainland Southeast Asia during the terminal Pleistocene and Holocene. Such an experimental design provides a platform for flintknapping experiments to avoid the inferential problems associated with replication and formal analogy, and move towards the construction of verifiable referential linkages on which solid relational analogies regarding past behavioral processes can be inferred.

Despite the inherent issues associated with replicative experiments, this kind of approach serves an important role as 'pilot' experimentation (Mathieu 2002) to identify and assess the importance of unknown variables and set of experiment protocols. Through this form of experiment, archaeologists can gain a general understanding of the procedure that is under question and how future research should be formulated. With this goal, the attempt to encompass all variables that would have been involved in past activities is not only justified but necessary (Comis 2010). At this phase, inferences are derived through formal analogy based on observations and experience of the experimenter. Although the security of the inference may be low, it illuminates the potential relationships among relevant variables.

In order for the potential relationships observed by pilot experiments to become firm referential linkages upon which relational analogy can be constructed, they need to be broken down further and independently verified by 'second generation' experiments. These experiments follow a clear protocol to ensure repeatability and allow quantifiable results (Mathieu 2002). At this stage, the emphasis is placed on the control of the

variable(s) in question rather than ecological validity. Therefore, experiments can be carried out with materials of a different nature to the original ones, depending on the hypothesis. In fact, because the variables have been evaluated before in the pilot experiment, they can now be tested in relation to each other in order to establish meaningful relationships to the studied phenomenon. Furthermore, as Domínguez-Rodrigo (2008) advocated, it is vital for studies to explicitly outline the assumptions and uniformitarian premises underlying experimental designs and interpretations in order for results to be effectively evaluated and compared. Through the falsification of test hypotheses, these sorts of experiment are able to establish concrete linkages of referential knowledge for drawing further archaeological inference.

Clearly, both kinds of experiments contribute to the study of past behavior, and it is important to acknowledge the limitations and potentials of each. However, the bulk of lithic inference construction should be based on second generation experiments with firm control over internal validity. That being said, the interaction between pilot and second generation experiment is not necessarily a one-way street but operates in a cyclical form of induction and deduction (Ingersoll and MacDonald 1977). Specifically, observations made from actualistic pilot data inductively generate ideas regarding the past phenomenon in question. These stipulated ideas are then deductively verified to see if the relationships are valid upon the ground of methodological uniformitarianism. Once a referential knowledge is established, new questions could once again arise and be formulated into testable relationships through actualistic studies. In addition, it is important for experimentally tested relationships to be related and compared to

archaeological data. After all, as Amick and colleagues (1989) related, "the ability to work back and forth between experimental work and the archaeological record is essential for learning about the past."

2.6 Conclusion

The issues regarding lithic experimentation and archaeological inference are not new. Indeed, numerous theoretical debates and discussions in the last several decades have been concerned with the nature of archaeological reasoning and how archaeologists should go about constructing sound inferences about the past (e.g., Fogelin 2007; Fritz and Plog 1970; Gibbon 1989). Stone artifact archaeology has come a long way since then, and employed increasingly sophisticated methods and theories for forwarding our understanding of prehistoric lithic artifacts. Recent years have witnessed experimentation playing a bigger part in lithic studies, and moving from the goal of simple reconstruction of reduction sequence to asking higher-level questions about hominin behavioral organization and evolution. Coupled with the growing role of archaeological knowledge in broader paleoanthropological discussions, the goal of this critical examination of lithic experimentation and analogic reasoning is to draw awareness to the importance of experimental design in lithic studies and its implication on the validity and security of the resulting inference.

It is hoped that as new territories of explanatory frameworks and analytical methods are explored and developed, lithic researchers can a maintain 'critical self-consciousness' (sensu Clarke 1973) of the field by carefully scrutinizing the rigor and

integrity in the way that they generate archaeological inference. This involves greater

transparency in the assumptions guiding hypothesis generation and the basic properties of

experimental design (Domínguez-Rodrigo 2008). Finally, as stated earlier, this essay

does not call for the rejection of replicative experimentation. Instead, both replicative and

controlled experimentation in the forms of pilot and second generation studies are

necessary for the investigation of the origin and use of lithic artifacts. In conjunction with

greater efforts to outline experimental design and its relationship to the overall

experimental process, lithic experiments can effectively provide constructive referential

linkages for building sound inferences of past behavior and adaptation based on

prehistoric stone artifacts.

CHAPTER 3: On the Utility and Economization of Unretouched Flakes: The Effects of Exterior Platform Angle and Platform Depth[2]

3.1 Abstract

In recent years, lithic studies have emphasized the role of technology in the overall adaptation of past societies to their environments, including the economization of lithic resources. This paper explores how particular characteristics of individual, unretouched flakes can be altered in ways that increase their economy, as reflected in the ratio of edge length to mass. Results of controlled laboratory experiments are presented that identify exterior platform angle and platform depth as being primary independent variables affecting this ratio. These relationships are then tested against a number of archaeological assemblages.

3.2 Introduction

The explanation of variability in lithic assemblages has been one of the fundamental goals in archaeology since the beginning of the discipline. Moving beyond the traditional cultural-historical paradigms based on artifact typology, archaeologists in the 1980s began to place emphasis on the organization of behavior that underlies overall variability in the archaeological record. In stone tool studies, this was reflected by a focus on the organization of technology (Bamforth 1986; Binford 1979, 1980; Potts 1991; Shott 1986; Torrence 1983). Relying on frameworks of behavioral ecology and optimal

---

[2] Author(s): Sam C. Lin, Zeljko Rezek, David R. Braun, and Harold L. Dibble. Author contributions: S.C.L., Z.R., H.L.D. designed research; S.C.L., Z.R., H.L.D. performed research; D.R.B. contributed data; S.C.L., Z.R., H.L.D. analyzed data; S.C.L., Z.R., H.L.D. wrote the paper.

foraging theory, archaeologists shifted attention from describing recurring tool forms to investigating the behavioral context of raw material acquisition, stone tool manufacture, use, transport, and discard. One of the main themes of these studies is the emphasis on behavioral strategies and their adaptive relationships with environmental constraints. In this paper a new dimension of variability with implications for adaptive significance will be explored, namely the production of flakes with a higher ratio of cutting edge to mass. Based on a series of controlled experiments (Dibble and Rezek 2009), it will be shown here that this ratio can be affected by certain parameters (specifically platform depth and exterior platform angle) that are directly under the control of a knapper, and the archaeological relevance of the interaction between these two variables is confirmed with reference to a large number of archaeological lithic assemblages.

In the framework of technological organization, lithic variability arises as a result of the dynamic interaction between behavioral strategies and the broader environment. These strategies "guide" the technological component of human behavior in relation to resource distribution, cost/benefit of time and energy, and risk management (Andrefsky 2009; Nelson 1991). The framework of technological organization is commonly conceptualized as different levels of strategic behavior organized hierarchically (Carr and Bradbury 2011; Nelson 1991). Artifact design and activity distribution are influenced by strategies such as scheduling and mobility, which are in turn structured by higher order economic and social strategies. The economic relationship between resources and the cost/benefit of energy has been argued to play a large part in the way these strategies are inferred. For example, the degree of time and energy invested in the production and use

of stone tools is seen as a technological behavior associated with risk management in relation to resource procurement (Bamforth and Bleed 1997; Bousman 1993, 2005; Elston 1990; Torrence 1983, 1989). And as resource distribution varies in both quality and quantity across the landscape, the economic relationship involved in resource procurement would have played a significant role in the organization of hunter-gatherer technology (Bousman 1993; Smith 1979; Winterhalder 2001). The importance of the design of lithic toolkits in relation to environmental conditions and group mobility is seen as being especially important, as the design of stone tools is a reflection of their immediate and future use (Bousman 1993; Jeske 1989; Kuhn 1994, 1995; Torrence 1983, 1989).

Thus far, much of the attention for studying the economic structure of lithic technology has been on retouched implements. Several discussions (e.g., Andrefsky, 1994; Bamforth, 1986, 1990; Bleed, 1986; Kelly & Todd, 1988; Kelly, 1988; Parry & Kelly, 1987; Shott, 1986) have focused on the design properties of mobile toolkits that facilitate different aspects of technological organization. These criteria include reliability, maintainability, transportability, risk management, time-stress, utility, and use life. Accompanied by an increasingly sophisticated methodology (e.g., Andrefsky, 2006; Clarkson, 2002; Iovita & McPherron, 2011; Iovita, 2011; Kuhn, 1990), various characteristics of retouched edges became a central focus for lithic analysts because it was felt that they provide observable and quantifiable units of intentional tool modification. The recognition of the effects of resharpening has also allowed the assessment of the extent of reuse and recycling among these formal tools as economic

strategies to buffer against the supply and demand of available toolstone. Thus, studies of retouched edges aim to interpret the extent of tool reuse and maintenance (e.g., Bamforth, 1986; Dibble, 1984, 1995b; Frison, 1968; Kuhn, 1992, 1995; Shott, 1989) and its effect on tool form and degree of curation (Andrefsky 2006, 2009; Shott and Sillitoe 2004, 2005; Shott 1996a).

Unretouched flakes, however, receive relatively less attention despite their clear significance in the organization of hunter-gatherer technology, as shown especially in ethnographic studies (Binford & O'Connell, 1984; Hayden, 1979b; Shott & Sillitoe, 2005; Sillitoe & Hardy, 2003; White & Thomas, 1972; White, 1967; see Holdaway & Douglass, 2012 for review), and by the fact that they usually represent the largest component of lithic assemblages. It has also been shown that flakes are deliberately produced and used in an unretouched state (Dibble and McPherron 2006). The fact is, however, that they are most often regarded as byproducts of particular core reduction strategies, or in terms of the selection of blanks for making retouched "tools." As a result, they are often not considered to be a central element of models concerning technological organization and optimization. While some measures involving unretouched flakes, such as blank-to-core or tool-to-flake ratios, are seen as reflecting reduction intensity and artifact recycling (Dibble 1995b), methods for quantifying the specific attributes of unretouched flakes that reflect efficiency, or relative utility, have yet to be fully developed in stone artifact studies.

This brings up the question: where is the utility in unretouched flakes? While characteristics such as size and durability clearly played significant roles in the production and selection of lithic artifacts (Braun et al. 2009b; Key and Lycett 2011; Prasciunas 2007), ethnographic data has repeatedly demonstrated the importance of a flake's usable sharp edge. For example, in their review of ethnographic accounts on stone artifact selection, Holdaway and Douglass (2012) argue that unretouched flakes constitute the fundamental functioning unit within lithic technologies and that the overriding purpose of many, if not most, lithic reduction technologies is simply to produce usable edges (also see Douglass, 2010). Kuhn (1994, 1995) and Morrow (1996; see also Roth and Dibble 1998) have discussed the expression of flake utility in terms of the amount of cutting edge vs. flake volume (weight) (also see Leroi-Gourhan 1964). This view underlies Shott and Sillitoe's (2005:657) argument that utility represents "the amount of use that a tool can supply in time, tasks performed, or other measures of use," and that utility can be realized from the artifact volume through edge rejuvenation and resharpening. In a sense, the amount of usable edge and the total volume of the artifact represent a trade-off between utilities that are immediately available and those that are potentially extractable in the future.

Momentarily putting aside its potential utility through subsequent (re-)modification, if an unretouched flake's utility, or value, is thus defined as the amount of cutting edge it provides, then any measure of economization for such flakes should be based on the amount of cutting edge in relation to the amount of material used. In other words, and to the extent that flake shape is constrained because of different raw material

69

properties or different uses, flakes with longer edges and less mass overall can be seen as being more economical, while flakes with shorter edges and more mass can be considered less so.

It does appear that certain core reduction technologies vary in the number of flakes produced and thus also in the total amount of cutting edge produced for a given volume of raw material. For instance, Levallois technology has been argued to be an efficient reduction strategy in both maximizing the amount of usable edge produced and minimizing waste of core preparation (Brantingham and Kuhn 2001; Dibble and Bar-Yosef 1995; c.f. Sandgathe 2004). Related to this, it has also been argued that the reduction of cores through what is known as "classic" linéal Levallois (Boëda 1994, 1995) results in flakes that are thinner and with mass that is more evenly distributed across their cross section (Eren and Lycett 2012; Van Peer 1992), which gives them a more viable working edge. Likewise, the systematic production of blades has been viewed as an example of enhanced technological efficiency for higher rates of flake production, as well as the large increase in the total length of cutting edge per volume of stone (Bar-Yosef and Kuhn 1999; Mackay 2008; c.f. Eren et al. 2008). However, these approaches largely remain restricted to categorical distinctions between technologies (e.g., Levallois vs. discoidal vs. blade) rather than relying on explicit measures of continuous variation in flake properties. An exception is Tactikos' (2003) quantitative assessment on the temporal trends of flake-edge-to-mass ratio across the Paleolithic and Mesolithic based on archaeological data and samples generated by replicative experiments.

Using experimental and archaeological data, it has been suggested earlier (Dibble 1997) that it is possible to control flake size and morphology through manipulation of certain platform attributes under the direct control of a knapper. Specifically, these platform variables are exterior platform angle and platform depth. What we will demonstrate here is that these same variables directly affect the length of cutting edge in relation to the mass of individual flakes.

3.3 Increasing the Ratio of Flake Edge to Mass

There are a number of flake attributes that directly affect edge length and mass. The first three to be discussed are based on solid geometry, while the last two take into account characteristics specifically relevant to stone flakes.
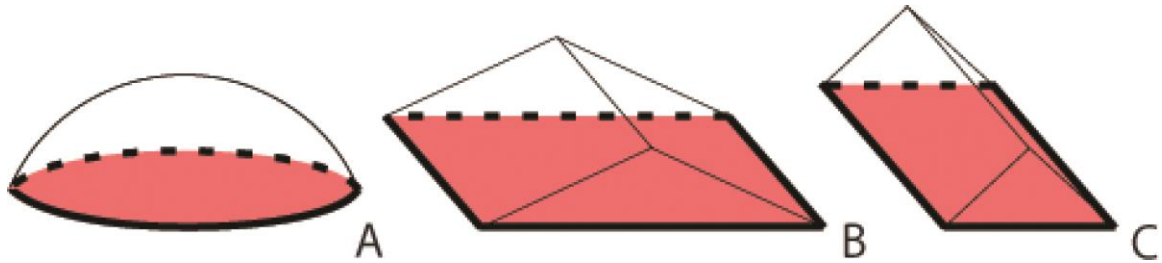
Figure 1 – Three geometric models of flake shape: (a) a sphere intersected by a plane; (b) a triangular prism with a square base (i.e., length = width); (c) a triangular prism with a rectangular base (i.e., length = 4 * width).

Figure 1 presents three geometric models that resemble flakes: a sphere intersected by a plane (Figure 1A) and two triangular prisms, one whose base length and width are equal (Figure 1B) and one whose length is four times its width (Figure 1C). Table 1 presents the numerical results of the first three strategies to be discussed.

Shape	Length (mm)	Width (mm)	Thickness (mm)	Ratio of Length to Width	Surface Area (mm²)	Surface Area/ Thickness	Edge Perimeter (mm)	Volume (mm³)
	2.22	2.22	0.49	1	3.88	8	6.98	1
Circular	4.44	4.44	0.97	1	15.51	15.99	13.96	8
	3.23	3.23	0.24	1	8.19	33.76	10.14	1
	2	2	0.5	1	4	8	8	1
Square	4	4	1	1	16	16	16	8
	2.83	2.83	0.25	1	8.01	32.04	11.32	1
	4	1	0.5	4	4	8	10	1
Rectangle	8	2	1	4	16	16	20	8
	5.65	1.41	0.25	4	7.98	31.92	14.13	1

Table 1 – Basic dimensions and dimension ratios of the three geometric models shown in Figure 1. Note: For each shape, the second row represents a doubling of all three dimensions given in the first row, while the third row decreases height in relation to the surface area of the base, while maintaining the same volume as the first row.

One strategy to increase the length of working edge, and perhaps the most obvious one, is to increase overall flake size. In Table 1, comparing the first and second rows for each geometric shape shows what happens when each of the three flake dimensions are doubled. While increasing overall flake size results in a doubling of edge length, this approach also has certain costs. The square-cube principle of proportional solids, sometimes referred to as allometry, results in an overall eight-fold increase in flake volume. In a very real sense, then, increasing flake size to obtain more edge is extremely uneconomical in itself, but it does remain an attractive option if the goal is to resharpen the flake repeatedly. A second strategy is to change the shape of the flake, and this can be done in two different ways. The first is to change the two-dimensional shape of the flake. Using the same three shape models, and with each shape having the same volume (of 1 unit³), it is clear from comparing the first rows of Table 1 that the three

72

different flake shapes have different perimeters, with circular ones having the lowest and the elongated rectangular prism having the highest. Thus, in knapping terms, increasing the ratio of length to width will result in flakes that have more usable edge per unit of volume. A second possible change in shape is to decrease flake thickness relative to surface area. As shown in Table 1, comparing the third row for each shape with its first row, decreasing thickness relative to surface area substantially increases edge length, even though volume is held constant.

While these simple geometric shapes can clarify certain relationships between edge length and flake volume, actual flakes are not quite as simple. In particular, two other considerations have to be taken into account. One of these is the size of the platform. While the models used above assume that the entire perimeter of a flake represents usable edge, in reality much of the proximal end of a flake is taken up by the striking platform. Not only does the platform represent an unusable edge, it also represents material that is taken from a core's striking platform, with the implication being that larger platforms will also diminish the use life of a core. Therefore, decreasing the size of the platform while maintaining edge length would not only conserve material that is part of the flake, but it would also help to maintain the core itself.

The final flake characteristic to be taken into account is the volume of the bulb of percussion. In most cases, the volume of material included in the bulb is essentially wasted mass, and in our experimental flakes, the volume of the bulb can exceed 20

73

percent of the overall flake volume. Therefore, decreasing bulb volume in relation to overall flake volume should be an economically desirable feature of flake production.

So, if the goal is to increase the absolute length of a flake's edge, one approach is (a) to increase the overall size of the flake, though this is in itself a relatively inefficient approach if the flake remains unretouched. It is much better, therefore, to change the shape of the flake so that it is either (b) more elongated or (c) has a reduced thickness relative to surface area. Either of these changes will help increase the edge length while limiting increases in overall volume. Other strategies that can be used in conjunction with the first two are (d) decrease the relative size of the platform, which both reduces the amount of unusable edge and maintains the core's striking platform surface, and (e) decrease the relative volume of the bulb of percussion, which is typically wasted material. While all of these attributes work together, in the following sections we will examine how a knapper can affect each of them when producing individual flakes.

3.4 Materials and Methods

Experimental data presented here were produced by the controlled experiment setup described in Dibble and Rezek (2009) and Rezek et al. (2011). It consists of a flaking apparatus with a core mount. Cores are securely clamped in the mount on the sides and the back, and the mount is adjustable for changes in the angle of the core platform surface relative to the angle at which the hammer strikes the platform (angle of blow) and also the distance from the point of percussion to the platform edge (platform depth, as defined by Dibble and Pelcin (1995); Dibble and Whittaker (1981); Figure 2).

The exterior surface of the core is completely exposed to prevent potential interference from the mount during flake formation. The hammer is made of steel and attached to a pneumatic cylinder; the tip of the hammer is shaped so that only the edge of it strikes the core. The extension of the pneumatic cylinder allows the hammer to hit the core platform and thus initiate the fracture that results in a flake. The cores themselves were manufactured from standard soda/lime glass with a semispherical surface morphology (Dibble and Rezek 2009), and thus are consistent in both size and shape. Because of the pneumatic cylinder, the applied force and velocity of each hammer displacement is uniform for every strike.
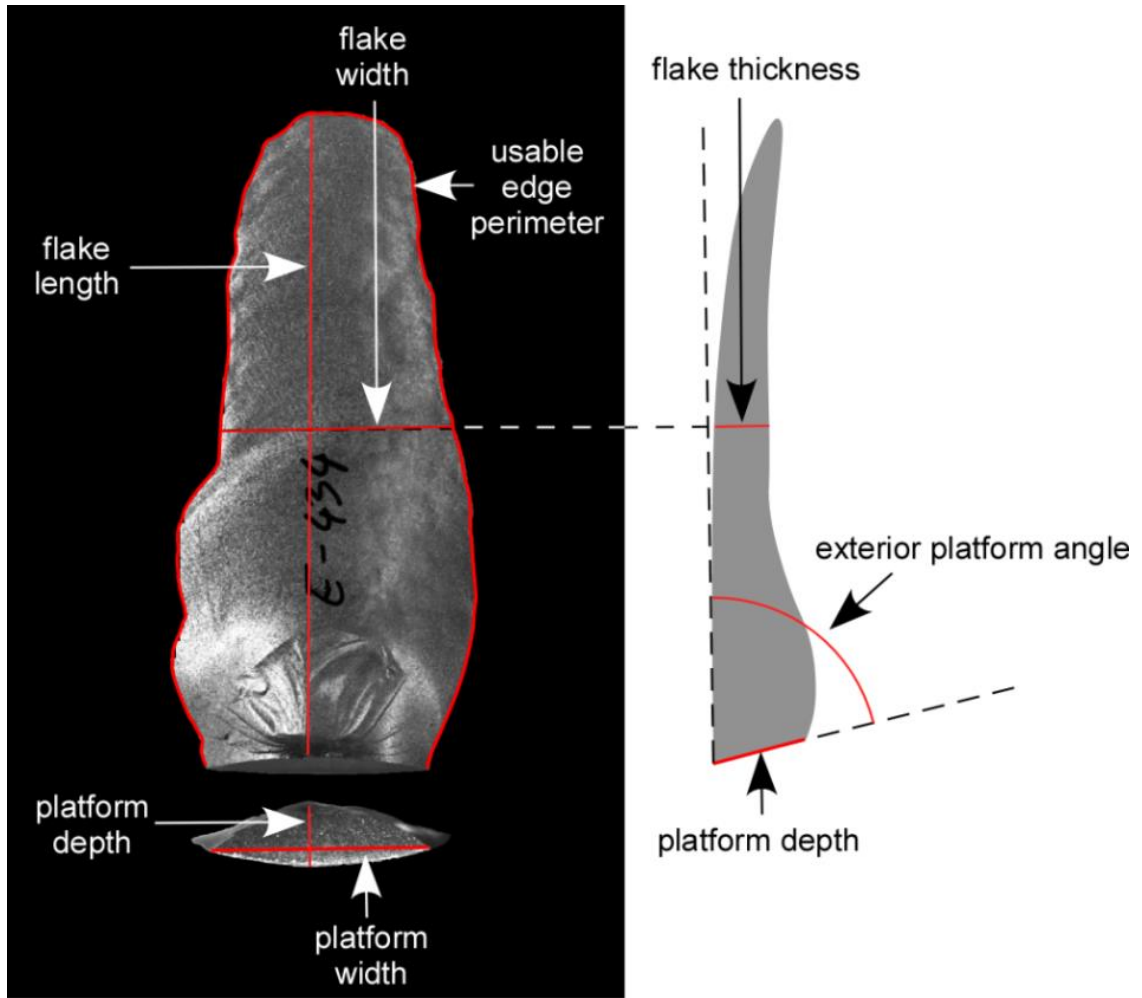
Figure 2 – Illustrations showing how basic flake measurements were taken. See text for more details.

In addition to the experimentally produced flakes, a large number of complete, unretouched flakes from various archaeological contexts were also included this study (Table 2). These data were collected by several individuals over the course of many years as part of their basic descriptions of these assemblages. While impossible to quantify at this stage, there is undoubtedly some degree of inter-observer error (reliability), although the manner in which all of the observations were taken has always been similar. It should be noted that in the subsequent analyses, no effort was made to control for various scar morphologies, raw materials, or other factors that may confound the results being investigated. In other words, all of the flakes greater than 2.5 cm in maximum dimension that had non-missing values for the specific variables being analyzed were included.

Assemblage	Location	Layer	Industry	N	Reference
1017	Nubia	N/A	Khormusan (MSA)	145	Marks 1968
FxJj1	Kenya	N/A	Oldowan	58	Isaac and Harris 1997
FxJj10	Kenya	N/A	Oldowan	157	Braun 2006; Isaac and Harris 1997
FxJj16	Kenya	N/A	Karari Industry	63	Harris and Isaac 1997
FxJj18G	Kenya	N/A	Karari Industry	95	Harris and Isaac 1997
FxJj20_M	Kenya	N/A	Karari Industry	113	Isaac and Harris 1997
FxJj63	Kenya	N/A	Acheulian	269	Harris and Isaac 1997
FxJj82	Kenya	N/A	Oldowan	87	Braun et al. 2009a
Brakfontein	S. Africa	N/A	LSA (Smithfield)	38	White and White 1964
Biache St-Vaast	N. France	IIA	Ferrassie Mousterian	1378	Dibble 1995a
Combe-Capelle	S. France	1D	Quina Mousterian	433	Dibble and Lenoir 1995
Combe-Capelle	S. France	1D1	Quina Mousterian	103	Dibble and Lenoir 1995
Combe-Capelle	S. France	1E	Quina Mousterian	365	Dibble and Lenoir 1995
Combe-Capelle	S. France	2A	Quina Mousterian	285	Dibble and Lenoir 1995
Combe-Capelle	S. France	2B	Quina Mousterian	114	Dibble and Lenoir 1995
Combe-Capelle	S. France	4A	Quina Mousterian	183	Dibble and Lenoir 1995
Combe-Capelle	S. France	4B	Quina Mousterian	125	Dibble and Lenoir 1995
Combe-Capelle	S. France	4C	Quina Mousterian	224	Dibble and Lenoir 1995
Pech IV	S. France	3a	MTA	1510	Turq et al. 2011
Pech IV	S. France	3b	MTA	2452	Turq et al. 2011
Pech IV	S. France	4a	Mousterian w Levallois	166	Turq et al. 2011
Pech IV	S. France	4c	Mousterian w Levallois	488	Turq et al. 2011
Pech IV	S. France	5a	Mousterian w Levallois	1211	Turq et al. 2011
Pech IV	S. France	5b	Mousterian w Levallois	382	Turq et al. 2011
Pech IV	S. France	6a	Mousterian w Levallois	1547	Turq et al. 2011
Pech IV	S. France	6b	Mousterian w Levallois	1311	Turq et al. 2011
Pech IV	S. France	8	Mousterian w Levallois	991	Dibble et al. 2009
Roc de Marsal	S. France	2	Quina Mousterian	345	Unpublished
Roc de Marsal	S. France	3	Quina Mousterian	104	Unpublished
Roc de Marsal	S. France	4	Quina Mousterian	1062	Unpublished
Roc de Marsal	S. France	5	Mousterian w Levallois	861	Unpublished
Roc de Marsal	S. France	6	Mousterian w Levallois	256	Unpublished
Roc de Marsal	S. France	7	Mousterian w Levallois	1327	Unpublished
Roc de Marsal	S. France	8	Mousterian w Levallois	1195	Unpublished
Roc de Marsal	S. France	9	Mousterian w Levallois	2835	Unpublished
Boker (D100A)	Levant (Israel)	Area A	Ahmarian	133	Marks 1977
Rosh Ein Mor	Levant (Israel)	Area A	Levantine Mousterian	213	Crew 1976
Rosh Ein Mor	Levant (Israel)	Area C	Levantine Mousterian	200	Crew 1976
Ein Aqev (D31)	Levant (Israel)	Lower	Levantine Aurignacian	128	Marks 1976
Ein Aqev (D31)	Levant (Israel)	Upper	Levantine Aurignacian	83	Marks 1976
Ein Aqev East (D34)	Levant (Israel)	All	Ahmarian	107	Ferring 1977
D40	Levant (Israel)	All	Levantine Mousterian	157	Munday 1976

Table 2 – Details on archaeological assemblages used in the analyses. *Note*: Only complete, unretouched flakes are included, as indicated by their respective N.

All independent variables, including angle of blow, hammer material, platform and core surface morphology, are held constant except for exterior platform angle (EPA) and platform depth (PD). EPA is measured at the intersection of the platform and the exterior core surface. To facilitate control over this variable, cores were designed so that there was no longitudinal curve along the exterior surface immediately behind the platform. For the experimental data used here, core EPA varies from 55 degrees to 95 degrees in 10-degree intervals. Different exterior platform angles are produced by making transverse cuts at the platform end with a diamond blade wet saw and measured with a goniometer. PD is measured from the point of percussion to the exterior edge of the platform. PD varies continuously between flakes. To minimize inter-observer error, four individual measurements of PD were taken on each flake separately and the mean was used for analysis. Dependent variables included in this study include flake weight, flake dimension, platform width, and the volume of the bulb of percussion. Flake weight is measured with an electronic scale to the nearest .1 gm. Flake dimensions consists of flake length (measured from the point of percussion to the most distal point on the flake), flake width (measured perpendicularly to flake length at the midpoint of the flake), and flake thickness (measured at the point of intersection of length and width). Platform width is measured perpendicularly to the axis of PD along the platform from one lateral edge to the other.

Four dependent variables were measured and calculated using a digitizing stylus Microscribe G2X and Rhinoceros software. These are flake edge length (measured to the nearest .01 mm), flake surface area (measured to the nearest .01 mm2) representing a

projected two-dimensional measure of the outlined flake surface, platform area

(measured to the nearest .01 mm2), and volume of the bulb (measured to the nearest .01

mm3) (Figure 3). For archaeological flakes, flake surface area was estimated by

multiplying flake length by flake width, while edge perimeter was measured as 2 x

(length + width). Bulb data are available for only a small portion of the archaeological

data and consist of measurements of the length of the bulb. Platform area was calculated

as Platform Width x Platform Depth. By comparing the results of using both ways to

record these variables on the experimental flakes, it is possible to show that the

differences are not significant (Table 3).

	r^2	p
Edge Length	.9475	<.000001
Surface Area	.9398	<.000001
Platform Area	.9974	<.000001
Bulb Volume	.5147	.00004

Table 3 – A test of the differences in which particular variables were recorded (using a Microscribe for the experimental flakes, linear measurements for archaeological flakes) by applying both techniques to the experimental flakes.
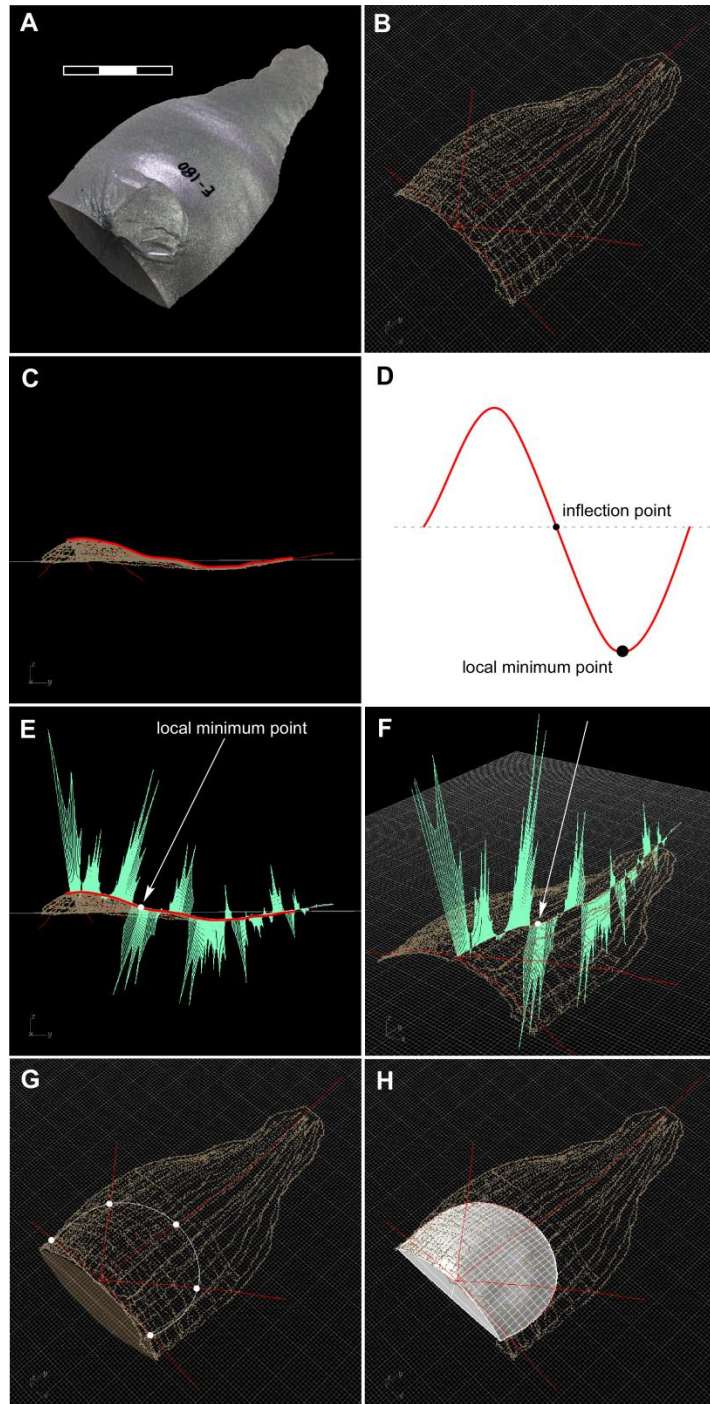
Figure 3 – Definition of bulb of percussion. After digitizing the interior surface of a flake (a) five curves were projected on that surface, emanating from the point of percussion and with 45° between each adjacent pair (b). Each of the five curves follows convexities and concavities present at the surface (c). With the aid of the curvature graph (represented by rays which are scaled to the degrees of convexity or concavity along the length of a curve in (e) and (f), the local minimum point of the first upward concavity from the point of percussion (d) is marked for each respective curve (g). These five points are then connected with another curve (g), which serves as a limit of the bulb in its definition (h).

3.5 Results

3.51 Increasing Overall Flake Size

Two independent experimental studies have shown previously that flake size is not a function of the amount of force applied to the core, even though larger (i.e., in terms of overall mass) flakes require more force to be successfully removed (Dibble and Pelcin 1995; Dibble and Rezek 2009). At the same time, the morphology of a core's flaking surface does not affect size either (Pelcin 1997a). However, and this has been confirmed repeatedly in many controlled experiments (Dibble and Pelcin 1995; Dibble and Rezek 2009; Dibble and Whittaker 1981; Pelcin 1997a; Speth 1981), flake size is undeniably a function of two variables: platform depth and exterior platform angle. Increasing either or both of those will result in larger flakes.

Figure 4 presents data obtained from recent controlled experiments (Dibble and Rezek 2009) concerning platform depth, exterior platform angle, and flake size. In this figure, note that platform depth is expressed as the cube, which renders the relationship linear due to the square-cube principle of proportional solids (platform depth increases in one dimension, while weight represents the combined effect of all three dimensions). There are two things to emphasize with these graphs. The first is that for each value of exterior platform angle, increasing the platform depth results in flakes with larger mass. Second, as the exterior platform angle increases, the relationship between flake mass and platform depth changes such that smaller increases in the latter result in even larger increases in the former. This is reflected by the slope of the regression equation, which shows that for lower values of exterior platform angle, flake weight increases relatively

slowly per unit increase of platform depth; but at higher values of exterior platform angle, flake weight increases much more rapidly for each unit that platform depth increases. Thus, platform depth and exterior platform angle each contribute to increasing the overall size (weight) of flakes.
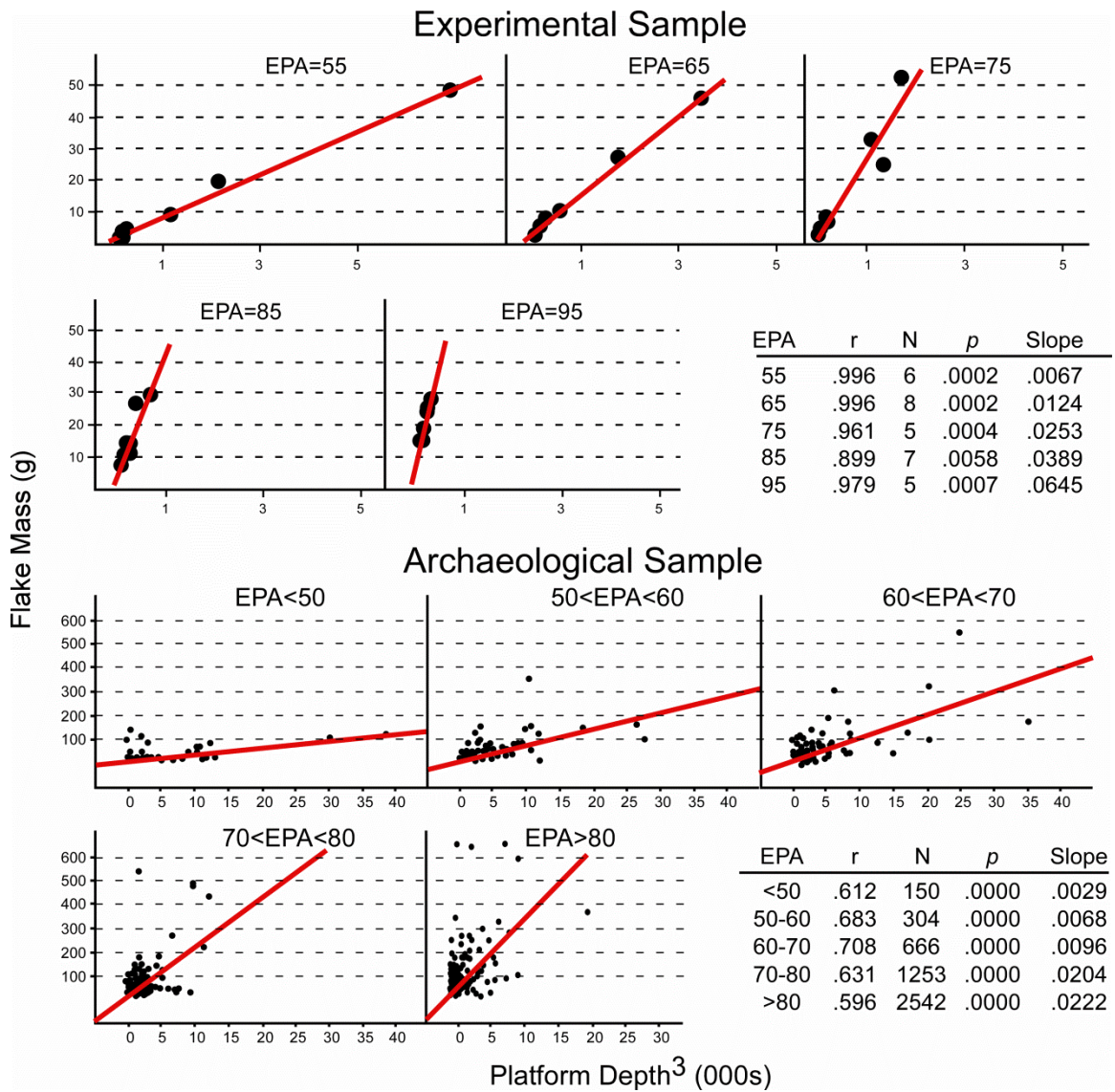
Figure 4 – Graphs of flake mass (vertical axes) and the cube of platform depth (horizontal axes) by increments of exterior platform angle for both the experimental and archaeological samples.

As shown in the lower portion of Figure 4, these exact same relationships can be seen in samples drawn from archaeological assemblages. This serves to confirm the results of the controlled experiment, and demonstrates as well that these relationships are quite fundamental in terms of basic flake production.

3.52 Changing Flake Shape

Increasing the Length to Width Ratio: Controlling the two-dimensional plan-view flake shape, in particular the ratio of length to width, is something that has been investigated over many years of replicative experiments, resulting in an almost axiomatic consensus that core surface morphology is the major independent variable that affects flake shape (Bar-Yosef and Kuhn 1999; Boëda 1986; Bordes 1961; Debénath and Dibble 1994; Inizan et al. 1995; Tixier et al. 1980). For example, it is most often believed that a core with parallel ridges will produce elongated blades, that a more rounded surface will produce circular flakes, and so forth. Recently, a controlled experiment was conducted focusing on core morphology and its effects on flake shape, and the results showed that while core surface morphology does influence flake shape, exterior platform angle has an even stronger influence (Rezek et al. 2011). When examined within a single core morphology (Figure 5), flakes made with higher exterior platform angles expressed higher ratios of length to width than those made with lower exterior platform angles. Moreover, these same results are obtained when using archaeological lithic assemblages without controlling for exterior scar morphology.

Figure 5 – Mean values of the ratio of flake length to width by increments of exterior platform angle for both the experimental and archaeological samples. For this and the following figures, F values were calculated with ANOVA.

Increasing the Flake Area to Thickness Ratio: As discussed above, the other change to flake shape that can influence the amount of usable edge relative to overall size is to decrease flake thickness relative to flake area. As already shown, platform depth is directly related to overall flake size, but as shown in Table 4, it affects each of the three linear dimensions of length, width, and thickness that all contribute to overall size. However, in both the experimental and archaeological data, the highest correlations between platform depth and flake dimensions are with flake thickness (Table 4). So, decreasing platform depth will decrease flake thickness more than the other two dimensions, but overall size will decrease as well. Then again, since overall size is also influenced by the exterior platform angle, decreasing platform depth while simultaneously increasing exterior platform angle will result in flakes that are still large in terms of surface area, but relatively thinner and lighter. Since both thickness and weight are reflecting the contribution of platform depth, dividing flake surface area by

85

either of these measures (while controlling for allometric effects) indirectly standardizes the effects of platform depth on these measures, thus allowing us to examine these variables with regard to exterior platform angle alone. In Figure 6, it is clear that in both the experimental and archaeological samples flake surface area, in relation to both flake thickness and weight, increases with increases in exterior platform angle.

| | Experimental Sample | | | Archaeological Sample | |
| | (N=35) | | | (N=10,886) | |
	r^2	p		r^2	p
Weight	.588	<.0001		.229	<.0001
Length	.17	.0112		.084	<.0001
Width	.704	<.0001		.231	<.0001
Thickness	.783	<.0001		.304	<.0001
Platform Width	.972	<.0001		.422	<.0001

Table 4 – R-squared and p values for correlations between platform depth and various flake measurements for both the experimental and archaeological samples.
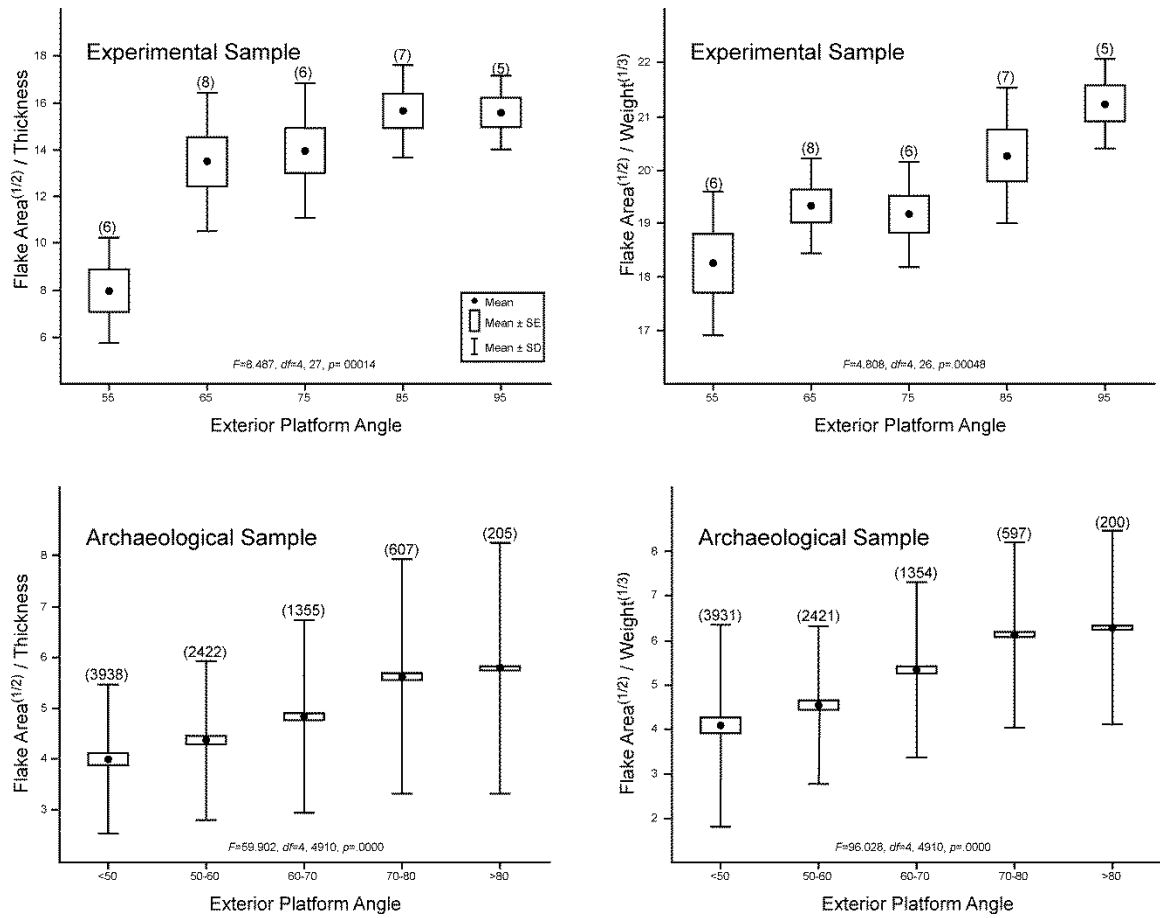
Figure 6 – Mean values of the ratios of flake area to both thickness and weight by increments of exterior platform angle for both the experimental and archaeological samples.

3.53 Reducing the Size of the Platform and Bulb of Percussion

So far the discussion has been on the size and shapes of flakes and the effects these have on increasing surface area in relation to thickness or weight. In this section, attention is turned to two other flake attributes, platform size and the volume of the bulb of percussion, which also relate to efficiency, since both represent potentially wasted material. The size of both of these attributes is also highly influenced through the combination of platform depth and exterior platform angle.

When the shape of the core behind the platform is held constant, platform depth has a strong and direct effect on platform width (Figure 7), though if the surface behind the platform becomes more convex or more concave, the ratio of platform width/platform depth either decreases or increases, respectively. However, detailed examination of those relationships, and strategies to control them, are beyond the scope of the present paper. What is of most interest here is not the shape of the platform, but its overall size, as represented by platform area, since it represents the portion of the core's platform surface that is removed along with the flake itself. All else being equal, flakes with smaller platform areas have less waste, and smaller flake platforms also minimize the effect of the flake removal on the core's striking surface. Since platform depth has such a strong effect on platform width, the strategy again is to decrease platform depth (which decreases platform width and, therefore, platform area as well), while maintaining flake size by increasing the exterior platform angle. The same strategy applies to decreasing the volume in the bulb of percussion.
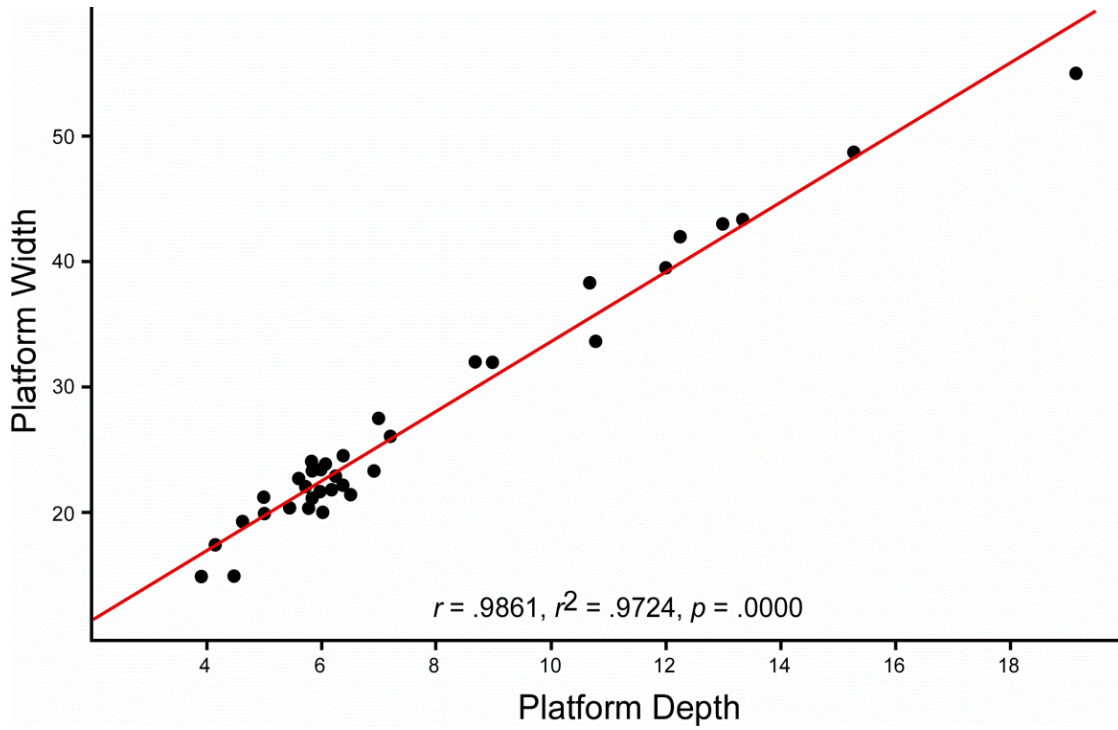
Figure 7 – Correlation of platform width and platform depth in experimental sample with constant core morphology.

Platform area is especially important in terms of maintaining the core surface area, but platform width alone, on the other hand, represents a portion of a flake's edge that is usually not useful. In examining the effects of platform depth and exterior platform angle on platform width, it is best to break down the overall edge perimeter into that portion represented by platform width and the other portion that is usable (of course, assuming no natural backing on the flakes). For both samples, then, the ratio that is most relevant is the percentage of usable edge, which is calculated as (total flake perimeter— platform width) / total flake perimeter. As shown in Figure 8, there is a clear relationship between exterior platform angle and this ratio in both the experimental and archaeological samples.
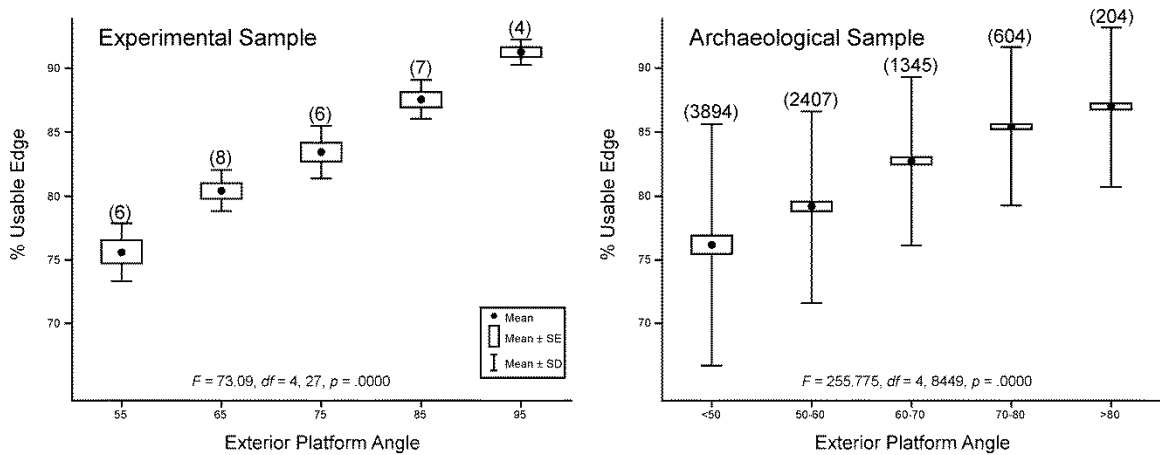
Figure 8 – Mean values of the percentage of usable edge (see text) by increments of exterior platform angle for both the experimental and archaeological samples.

Finally, equally clear effects are seen between exterior platform angle and the size of the bulb of percussion (Figure 9) in that higher exterior platform angles result in smaller bulbs of percussion.
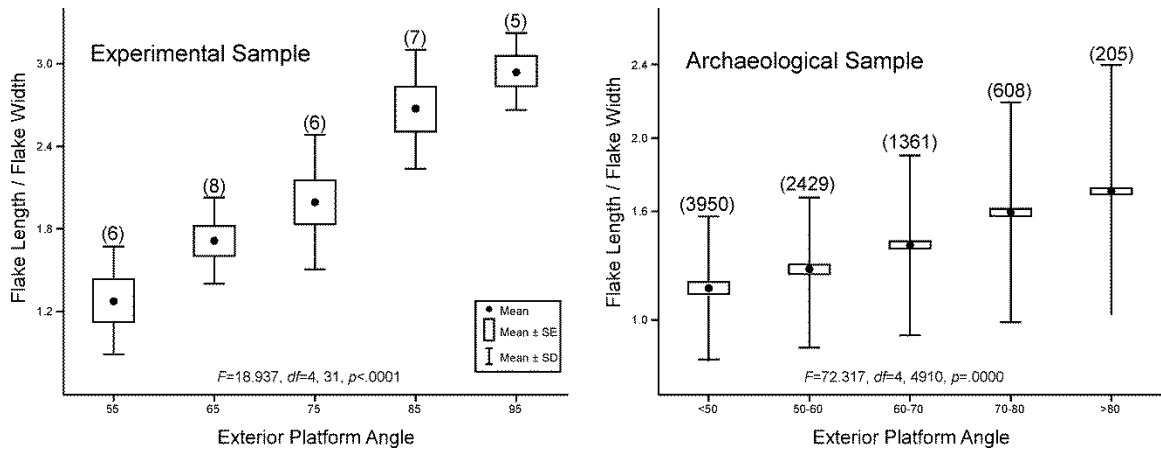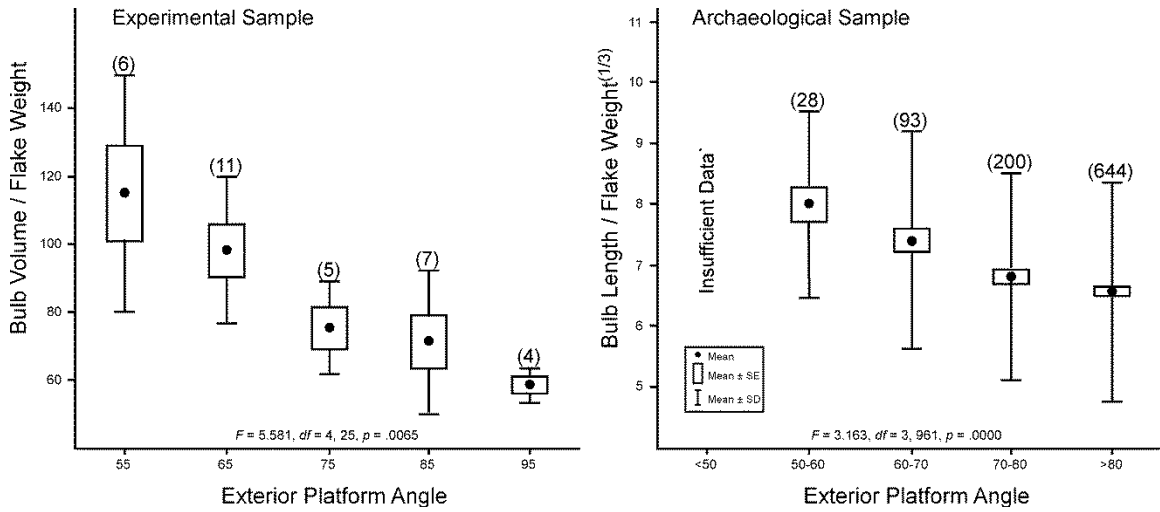

Figure 9 – Mean values of the ratios of bulb volume to flake weight by increments of exterior platform angle for both the experimental and archaeological samples.

So far we have examined, in turn, several variables that contribute to increasing the edge length in relation to mass, and in each case exterior platform angle plays a significant role. It should also be true, therefore, that the primary ratio of interest— edge length to mass— is similarly affected by this independent variable. Figure 10 shows that this is the case.
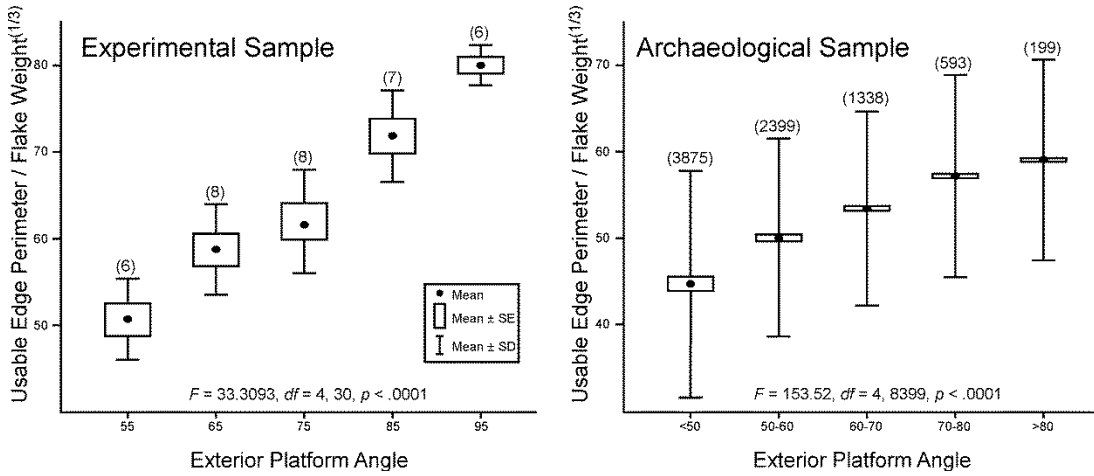
Figure 10 – Mean values of the ratios of usable edge perimeter (see text) to the cube root of flake weight by increments of exterior platform angle for both the experimental and archaeological samples.

It should be clear at this point that through an interplay of both platform depth and exterior platform angle— two variables that are under the direct control of the knapper— it is possible to effectively increase the ratio of usable edge to flake mass. The above results are summarized in Figure 11, which illustrates the various effects of these two variables as two diagonal axes. First, overall flake size is determined by increasing or decreasing both of these variables at the same time, resulting in larger or smaller flakes, respectively. If, however, one of these variables is decreased while the other is increased, the result is a number of changes in a flake's morphology that, when analyzed either individually or together, alter measures of the overall economy of the flake.
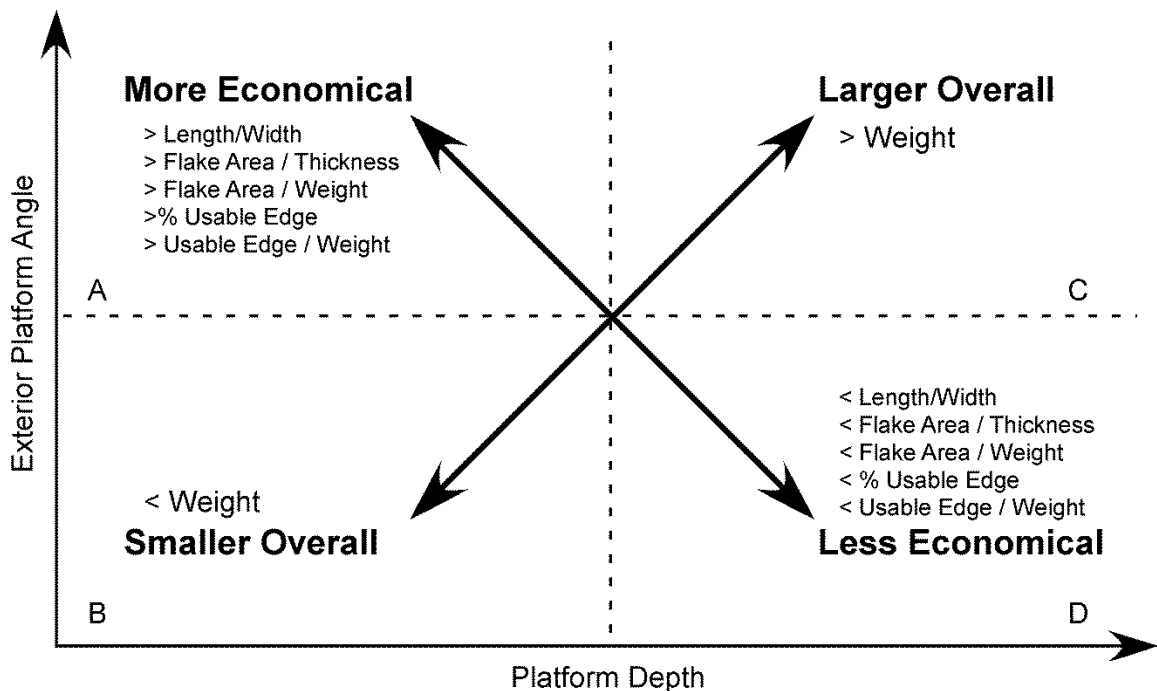
Figure 11 – Schematic diagram illustrating the relative effects of exterior platform angle and platform depth on variables reflecting more or less economical flakes and smaller vs. larger overall size. The dashed lines reflect the mean values of exterior platform angle and platform depth for the archaeological assemblages used here, thus creating the four groups of assemblages that are compared in Table 4.

93

While the preceding analyses were based on individual flakes, it is also possible to test these effects by comparing mean values of the relevant variables computed for whole assemblages of flakes. First, the mean values of exterior platform angle and platform depth are computed for each archaeological assemblage, and then the average (for all assemblages) of those means (= 76.8 and 6.9 for exterior platform angle and platform depth, respectively) is used to divide the assemblages into four groups: those with higher or lower than average values of exterior platform angle, and those with higher or lower than average values of platform depth. Then it is possible to compare the average values of each of the groups for the dependent variables analyzed thus far (with the exception of bulb length, which was recorded on very few of the assemblages used here). Thus, and referring again to Figure 11, we can compare the average assemblage values of weight between groups B and C (the opposite extremes along the size axis) and the average values of the other variables between groups A and D along the economical axis. As shown in Table 5, all of these comparisons fall in the direction predicted, and all are significant. Thus, all of the relationships discussed above are clearly apparent even at the level of assemblages, rather than individual flakes. Since the relationships are the same in both cases, it shows that the effects crosscut major technologies and exist, therefore, at the very fundamental level of the mechanics of chipped stone.

	High EPA Low PD (Group A) Mean	Low EPA High PD (Group D) Mean	T	df	p
Length / Width	1.74	1.40	4.55	23	.000143
Flake Area / Thickness	5.62	3.89	5.97	23	.000004
Flake Area / Weight	16.20	14.68	6.72	23	.000001
% Usable Edge	.86	.84	3.85	23	.000822
Usable Edge / Weight	58.99	50.25	9.69	23	.000000

	High EPA High PD (Group C) Mean	Low EPA Low PD (Group B) Mean	T	df	p
Weight	47.59	10.19	5.44	15	.000068

Table 5 – Statistical comparisons of the groups defined in Figure 11 on the basis of variables that reflect economization (upper) and overall size (lower).

3.6 Assessing Economic Behavior in Archaeological Assemblages

What has been demonstrated so far is that the interaction of EPA and PD significantly affects flake morphology. These relationships are not only seen under controlled experimental conditions, but are also apparent in archaeological samples. While we have argued that the kinds of changes in flake morphology described here relate to varying degrees of economization— and these arguments are based on simple geometric principles— it would be more satisfying if it were possible to test whether or not hominin populations actually employed such strategies in increasing the length of working edge under particular conditions, such as in situations of raw material scarcity. However, our inability to control, characterize, or quantify such external factors makes such a test impossible, at least with the methods currently at our disposal.

Even so, it is possible to show that there is variation in the two independent variables, and in the concomitant effects that they have on flake morphology, among archaeological assemblages. For example, Figure 12 plots a series of French Mousterian assemblages from three sites: Pech de l'Azé IV (Turq et al. 2011), Combe-Capelle Bas (Dibble and Lenoir 1995), and Roc de Marsal. In traditional Bordian systematics based on relative proportions of different tools (Debénath and Dibble 1994), these assemblages represent several different variants. In this figure, however, two groups emerge: one contains all of the assemblages from the various sites that are identified as belonging to the Quina Mousterian, while the other represents a number of non-Quina variants. Based on the above analysis, the Quina group would appear to be the least economical group, and the non-Quina group would appear to be the more economical one.
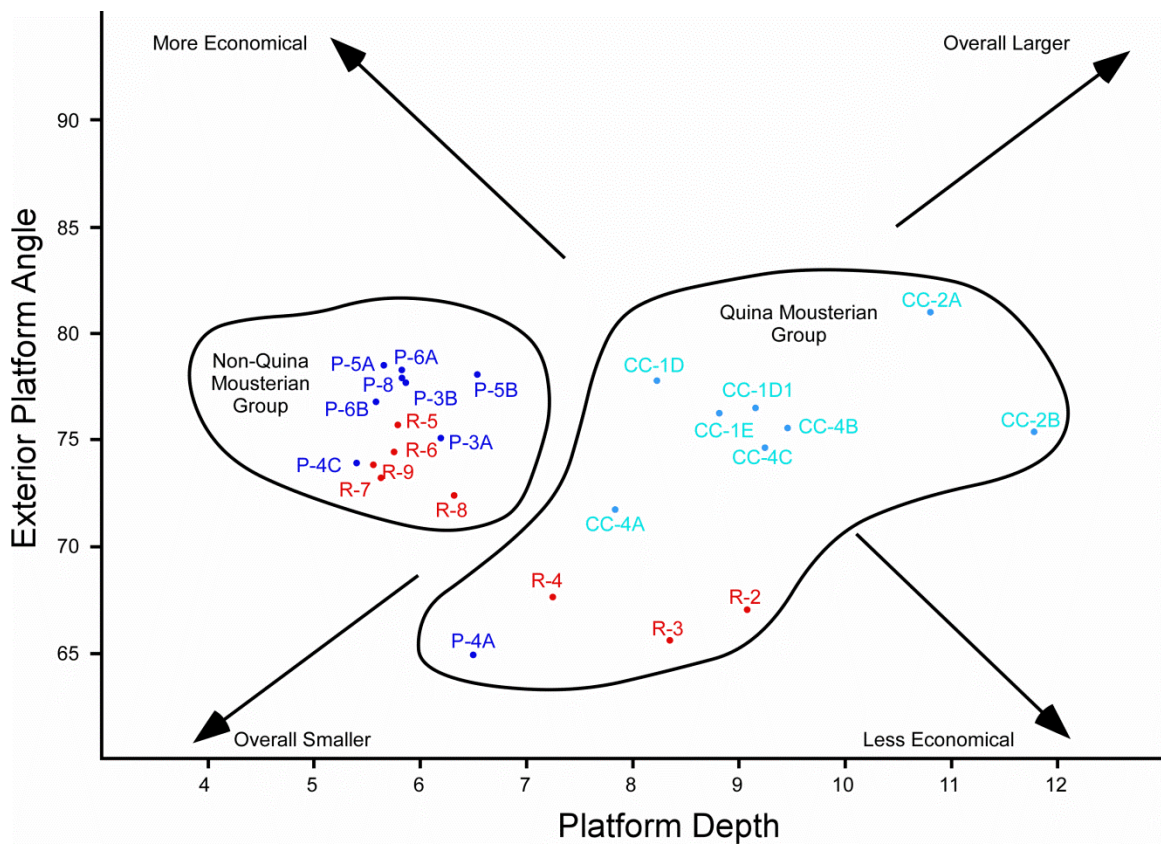
Figure 12 – Various French Mousterian assemblages plotted according to their mean values of EPA and PD.

In fact, this dichotomous relationship within Mousterian assemblages is reflective of two distinct strategies of economization. One, evident in the non-Quina group, is the production of flakes, especially those produced through Levallois technology, that have higher edge to mass ratios (Figure 13). There are relatively fewer retouched pieces in these assemblages, and the degree of resharpening (as measured by those "tools" that reflect more or less resharpening episodes; Dibble 1995b) is also less. The other strategy, evident in the Quina group (Figure 14), emphasized the production of flakes having less

edge margin to mass, but these assemblages reflect a distinctly different approach to conservation of material. As has been demonstrated earlier (e.g., Rolland and Dibble 1990; Turq 1989, 1992), many characteristics of Quina Mousterian lithic assemblages reflect an emphasis on repeated tool resharpening. Such a strategy is based on the production of flakes with a suitable shape and volume to maximize the number of resharpening episodes. Undoubtedly, these two approaches are not mutually exclusive, since more economical flakes can themselves be resharpened. Nevertheless, they both show deliberate, though different, attempts to increase the efficiency of their lithic technologies, one through changing the morphology of the flakes to provide more usable edge per units of mass (the non-Quina group) and the other by maximizing the resharpening potential of the flakes. While it remains unclear as to why one strategy would be adopted over the other, clearly each group manipulated the two independent variables of EPA and PD to produce different flake morphologies, and each approach can be viewed as having different effects on economization. What this potentially suggests is that by the time of the French Mousterian, at least, hominins were aware of the need for more efficient products and developed varying strategies for achieving them.
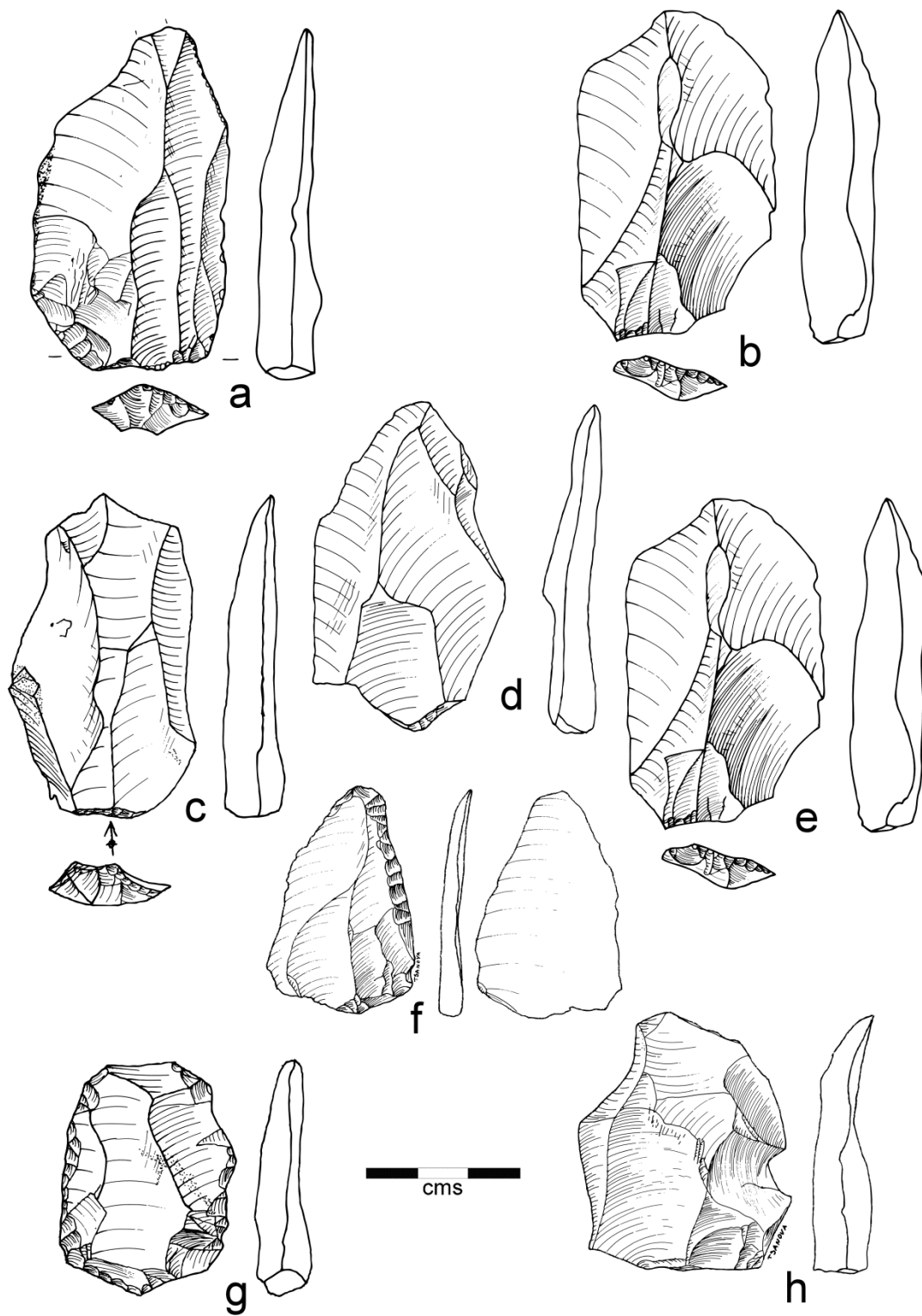
Figure 13 – Lithic artifacts from various non-Quina assemblages plotted in Figure 12. (a), (b), (e) Pech IV, Level 5A; (c), (d) Roc de Marsal, Level 9; (f) Pech IV, Level 6a; (g) Roc de Marsal, Level 5; (h) Pech IV, Level 3B.
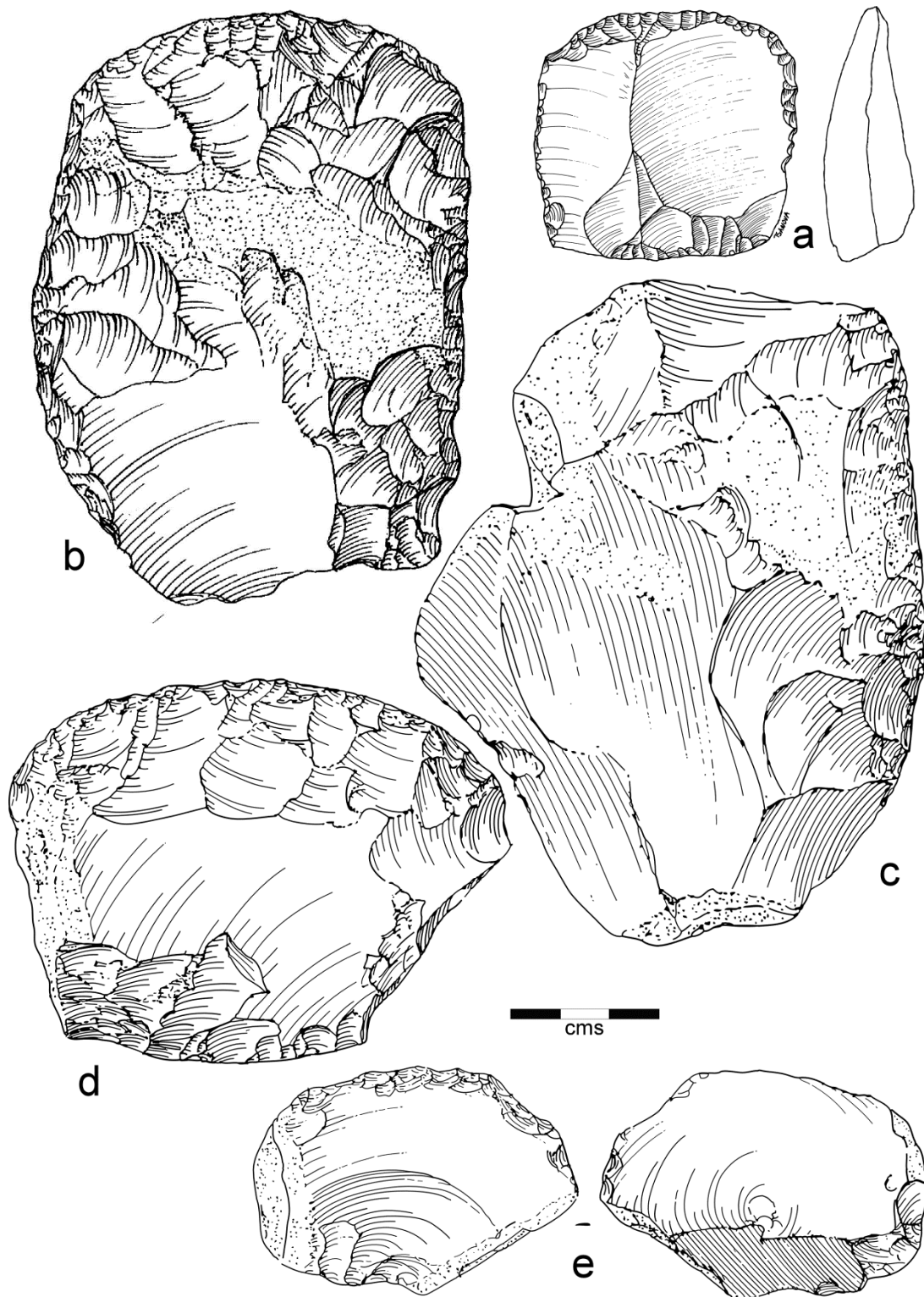
Figure 14 – Lithic artifacts from various Quina assemblages plotted in Figure 12. (a) Pech IV, Level 4A; (b) Combe-Capelle Bas, Level I-2B; (c) Combe-Capelle Bas, Level I-2A; (d), (e) Combe-Capelle Bas, Level I-1E.

3.7 Discussion and Conclusions

For many years it has been emphasized that economization of resources constitutes an important explanation for many aspects of lithic variability (Odell 1996). As discussed in the opening of this paper, such explanations have largely focused on the design or maintenance of retouched pieces or on major technologies. Until now there has been little discussion of how economization works at the level of the production of individual flakes. To a large extent, this lag is a result of the lack of solid middle range theory (Binford 1977b) concerning the most fundamental aspects of flake production and the means by which knappers can alter specific characteristics of the flakes they produce. Highly controlled experiments are now beginning to contribute significantly to that body of theory, and to a large extent, such experiments allow us to completely turn around the scientific process through which we can develop new models of lithic variability. Instead of drawing on the archaeological record to generate hypotheses that are then "tested" through replicative studies, it is now possible to identify specific cause-and-effect relationships in the lab and then proceed to test them with the archaeological record.

This latter process is the one that has been used here. This study began with the assumptions that (a) people relying on stone tool technologies use unretouched flakes; and (b) that the cutting edge represents one fundamental utility of an unretouched flake. Thus, one way to economize on lithic resources would be to increase the amount of cutting edge in relation to the amount of mass contained in a flake. There are then two levels of understanding as to how this can be achieved. The first level is theoretical. By drawing on models of solid geometry and particular attributes of chipped stone flakes, it

is possible to identify a number of characteristics— including overall shape, platform width, and bulb size— that can contribute to increasing edge length to mass. The second level is more practical— determining how those characteristics can be altered by a knapper— and it was facilitated by the experimental design. Based on these results, the variables of platform depth and exterior platform angle— both of which are directly under the control of the knapper— emerged as the two most important. The overall size of the flake is controlled by both of these variables working synchronously— increasing both platform depth and exterior platform angle results in larger flakes, and reducing both of them results in smaller flakes. However, in terms of all of the shape and other flake characteristics contributing to the ratio of edge length to mass, it is the adjusting of one of these independent variables in relation to the other that makes the biggest difference. Increasing the exterior platform angle while decreasing platform depth results in more economical flakes; decreasing the former and increasing the latter results in less economical flakes. The final step was then to test these relationships with flakes recovered from the archaeological record. This was also done on two levels. The first was at the level of individual flakes (combining flakes from several different industries and technologies), while the second was at the level of whole assemblages— comparing mean values of the dependent variables from different assemblages that differed in terms of the mean values of the two independent variables. Finally, an example was presented showing that in the French Mousterian these two independent variables were manipulated in two different ways, each contributing to different approaches to economization.

It should be noted that altering exterior platform angle and platform depth are not the only strategies that can lead to increased ratios of edge length to mass. For example, an increase in the length to width ratio can also be facilitated through core preparation, though, as shown earlier (Rezek et al. 2011), core morphology, at least in terms of configuration of core surface ridges, plays a more minor role relative to the two independent variables discussed here. It is also possible, though not yet demonstrated under controlled conditions, that the use of an indirect "punch" technique may contribute to producing smaller platforms. And, as discussed above, certain patterns of core reduction may also contribute to overall economization. Our point here, however, is that such effects may also be achieved at the level of individual flake production.

What is interesting about the archaeological assemblage data is that assemblages vary significantly in terms of all of the attributes discussed here, and thus also in terms of the overall economization of their unretouched flakes. Moreover, the temporal range of these assemblages spans almost 2 million years of technological variation— from the Oldowan through the Later Stone Age. This suggests in itself that these are fundamental properties of flake production and not just the result of particular technological patterns. At this point it is unknown what the underlying factors are that gave rise to that variability, though it is likely that access to suitable raw material is a significant one.

This also has implications for concepts such as curation and expediency. The decision to abandon or retain certain materials is likely to be based on considerations of utility (Shott 1996a), transportability (Nelson 1991), usable edge proportion (Kuhn

1994), and quality (Brantingham et al. 2000; Roth and Dibble 1998). This economic relationship between mobility and technological organization lies at the center of the curation concept proposed by Binford (1973, 1977a). Despite debate regarding the nature and use of the term (see Andrefsky 2009 for review), curation captures the economic aspect of behavior in the production and maintenance of stone artifacts for maximizing or prolonging their utility. If the utility of unretouched flakes is in the amount of cutting edge they provide, then the results presented here clearly show how utility can be altered in terms of the economization of raw materials. As such, curation behavior can also be seen in the economic production of unretouched flakes, even in the absence of the production of more formal "tools" (Douglass 2010). In current approaches, the production of unretouched flakes alone would be taken by many as indicative of some sort of "expedient" lithic technology (Bamforth 1986; Parry and Kelly 1987), even though that might not be the case. This points to an inadequacy of an overly simplified theoretical dichotomy between "curated" (or "formal") and "expedient" artifacts, and it also suggests that we should avoid focusing exclusively on retouched artifacts when addressing issues of curation (Holdaway and Douglass 2012).

In this same vein, however, it is also important to recognize that whatever the economy and utility apparent in the unretouched flakes, it does not necessarily adequately reflect the complexity or efficiency of the overall underlying technology. In other words, assemblages with less economical flakes in terms of cutting edge length to mass ratio, such as those seen above in the Quina Mousterian, do not necessarily translate to less efficient lithic strategies than those that produced thinner and broader flakes (Carr and

Bradbury 2001). Hunter-gatherer technological organization involves a complex mixture of activities, strategies, and decision-making (Bamforth 1991; Torrence 2001). The same goal can be achieved or similar problems resolved through many different combinations of different actions. What this means is that technology is multidimensional (Chatters 1987) and strategies that underlie the production of individual flakes represent just one of those dimensions. It is an important dimension, however, and one that should be taken into account more fully in future studies of lithic technological variability. Finally, it is not the intent of this paper to argue that hominins were always under pressure to economize their raw materials. Rather, our purpose here is to present a "proof of concept," namely, to describe how an individual knapper can manipulate certain platform variables to change the overall morphology of flakes. Moreover, such changes can likely be interpreted in economic terms rather than just simple function vs. style alternatives. Ultimately, it will be essential to test these patterns with independent data, especially raw material quantity and accessibility, but perhaps also group mobility and other relevant factors. It is important to keep in mind, however, that there are undoubtedly many ways to economize raw materials, and just as important, there are undoubtedly many situations in which such economization is not necessary. Our point here, therefore, is to introduce some of the ways in which economization can take place on the fundamental level of individual flake production.

CHAPTER 4: Establishing Statistical Confidence in Cortex Ratios between Lithic Assemblages: A Case Study of the Middle Paleolithic of southwestern France[3]

4.1 Abstract

Recent studies have demonstrated the usefulness of the Cortex Ratio for quantifying the cortex composition in lithic assemblages and as a viable index of prehistoric artifact transport. Yet, the lack of means for assigning statistical confidence to archaeologically observed Cortex Ratios inhibits the approach's utility for objective comparisons and interpretation. We derive statistical confidence for archaeological Cortex Ratios through Monte Carlo and resampling techniques. Experimental data with known geometric properties and measured cortex values were employed as a reference for attaching a probability to an archaeological assemblage's Cortex Ratio. The method is demonstrated on assemblages from the Middle Paleolithic sites of Roc de Marsal, Pech de l'Azé IV, and Combe-Capelle Bas in southwestern France.

4.2 Introduction

In stone artifact archaeology, cortex is an attribute commonly used for assessing reduction intensity and sequence (Andrefsky 2005; Dibble et al. 2005), raw material exploitation and transportation (e.g., Reher 1991), site use (e.g., Roth & Dibble 1998) and mobility (e.g., Fernandes et al. 2008; Olszewski et al. 2010; Kuhn 1991, 2004). Because lithic technology is reductive in nature, the amount of cortex retained on stone artifacts is

---

directly correlated with the degree of nodule reduction (Dibble et al. 2005; Douglass et al. 2008). However, as Dibble et al. (2005:545) noted, it often remains unclear whether an assemblage has more or less cortex than expected given models of varying site use, curation, and technological organization. A large part of this uncertainty relates to the variability in the initial cortex abundance of lithic assemblages caused by differences in the size and shape of the cobbles from which artifacts were produced from. Recognizing this issue, Dibble et al. (2005) established an objective measure of the expected amount of cortex that should be observed for a given quantity of stone assuming when fully cortical nodules were reduced on site and all products of reduction remained on site. The approach is based on estimates of the geometric shape of unworked stone cobbles, the volume of the assemblage measured and the number of nodules worked.

This measure of expected cortex is compared to the total cortical surface present in the assemblage. The relationship between these two values, expressed by the Cortex Ratio as the amount of cortex observed in an assemblage versus the amount expected, thus provides a mean to determine whether all of the products resulting from nodule reduction are present at a location, or if some elements were either removed or added. If the Cortex Ratio is equal to 1, then it suggests that all of the knapped elements are present. If the ratio is less than or greater than 1, then it suggests that less or more cortex, respectively, is present than would be expected under the assumption of fully cortical nodules knapped in place without subsequent transport. While the concept is clear, further studies are needed to objectively interpret archaeological Cortex Ratios. Specifically, we would like to know how far the ratio has to deviate from a value of 1 to

indicate with confidence the effects of artifact transport. Likewise, we currently lack a

sound statistical basis with which to interpret variation in Cortex Ratio values for

different archaeological samples. The purpose of this paper is to investigate the use of

Monte Carlo sampling approaches to derive sampling distributions for the Cortex Ratio

that, in turn, will allow us to assign a probability for rejecting or accepting the null

hypothesis that differences between archaeological Cortex Ratios are due to sampling

error alone. We will then apply this method to the Cortex Ratios from several French

Middle Paleolithic assemblages.

4.3 Background

Aside from the original experiments by Dibble et al. (2005), the robustness of this

methodology has also been repeatedly verified by other experimental testing (Douglass

and Holdaway 2011; Douglass 2010; Douglass et al. 2008; Holdaway et al. 2008; Lin et

al. 2010; Parker 2011). Subsequent applications of this approach demonstrated its

feasibility for assessing the relative extent of artifact transport and, hence, the degree of

past mobility (Dibble et al. 2012; Douglass 2010; Douglass et al. 2008; Holdaway et al.

2010, 2012, 2013; Phillipps 2012). Differences in cortex composition among lithic

assemblages therefore provide an objective and quantitative way of comparing variation

in the patterns of past movement and technological behavior.

To date, the most thorough application of the cortex approach was by Douglass

(2010; also see Douglass et al. 2008) with the mid-to-late Holocene surface lithic

assemblage in western New South Wales, Australia. Douglass examined a sample of over

170,000 stone artifacts from four study locations, and one excavation, across the region. Cortex Ratios ranging from 0.16-0.68 indicated that cortex was consistently underrepresented at all locations. Because the lithic artifacts were almost exclusively produced from locally abundant stones, the underrepresented cortex most likely reflects repeated removal of the larger flakes, which also tend to be disproportionally cortical (Roth and Dibble 1998) from the sampling localities. This interpretation is supported by experimental simulations of artifact transport (Parker 2011), which suggested the cortex deficits can be accounted for by removals of up to 25% of the assemblage in large flakes. The spatial scale of the cortex pattern and its recurrence in a broad sample of assemblages further indicate high mobility where past populations travelled over wide territories, possibly taking advantage of ephemeral opportunities for occupation in an arid and extremely unpredictable landscape (Douglass 2010; Douglass et al. 2008; Holdaway and Allen 2013; Holdaway and Douglass 2012; Parker 2011).

Phillipps' (2012; also see Holdaway et al. 2010) study of the lithic assemblages at stratified Neolithic sites in the Fayum, Egypt, on the other hand, suggests that lithic raw materials were transported to the former lake margin. In these assemblages, Cortex Ratios range from 0.7 to 0.9, which suggest a considerable amount of cortex remained in the assemblages. Because workable stone does not occur naturally in the study area, cortical cobbles were likely transported as cores to the lake margin. This observation was supported by the higher frequencies of cores than the expected number of worked nodules in the assemblages. The higher Cortex Ratios approaching 1 indicate that, while cortical nodules were transported to the lake shore, reduced artifacts were not

109

consistently moved away from the area in sufficient distances, signaling that substantial

activities of the Neolithic populations occurred close to the lake edge.

More recent work examined the Cortex Ratios at the Middle Stone Age site of

Contrebandiers Cave, Morocco (Dibble et al. 2012). The industries at the site (so-called

Aterian and Maghrebian Mousterian), which are associated with North African modern

Homo sapiens in the Upper Pleistocene, are mainly made on quartzite that is available

200m from the site, along the coast (Bouzouggar 1997; Dibble et al. 2013). Applying the

Cortex Ratio to the quartzite artifacts, the Maghrebian Mousterian assemblage produced a

value of 0.7 while the four Aterian assemblages yielded a set of rather consistent values

in the 0.5 range. This general deficiency in cortex across the assemblages was taken to

reflect that local lithic materials were transported away from the site. Although it is true

that the same cortex pattern could be created by the import of decertified cores, this

scenario was argued to be unlikely due to the close proximity of the raw material sources.

This issue of equifinality could potentially be clarified by further simulation work, such

as the study by Parker (2011). The discrepancy between the ratio values of the two

industries at Contrebandiers Cave was interpreted to reflect differences in the mobility

and land use strategies employed.

These studies have all helped to demonstrate the effectiveness of the cortex

methodology in capturing the relative amount of cortex to volume of a given

archaeological sample. However, a ratio value is simply a number at this point, and it is

less clear how different Cortex Ratios can be compared objectively. This problem raises

two key issues. First, how do we assess whether the cortex composition of a given assemblage is different from that of a complete assemblage not influenced by artifact transport? That is, how can we determine with confidence that a ratio value above or below 1 does, indeed, indicate that transport has affected assemblage composition? For example, do the ratios of 0.7 to 0.9 observed by Phillips (2012) at the Fayum reflect real cortex deficits, or could they instead be due to sampling error? The second issue relates to the method of determining if Cortex Ratios between two different assemblages are indeed different from one another at a given level of statistical significance and thus reflect different patterns of production, selection, transport and discard. While the Aterian assemblages at Contrebandiers with ratios in the 0.5 range have less cortex relative to artifact volume than the Mousterian assemblage that has a ratio of 0.7, it is difficult to say immediately whether this difference is significant or, again, whether it is simply due to sampling error.

4.4 Materials and Methods

The establishment of statistical confidence of Cortex Ratios is investigated here through the use of sampling techniques on both experimental and archaeological data. The archaeological data used are from three Middle Paleolithic sites located in southwestern France. The rationale for the use of these sites is that they contain a range of assemblage sizes among stratigraphic layers that allows the assessment of sampling error in Cortex Ratios. Their spatial proximity and diachronic lithic sequence spanning the late Pleistocene across different Mousterian industries also offers the potential for comparing different Cortex Ratio values with existing models of Neanderthal mobility in

Western Europe (e.g., Delagnes and Rendu 2011). Because the cortex quantification method is partly contingent on the contextual information of the study assemblages, the archaeological data are outlined first below and followed by descriptions of the methodological procedure and the experimental data employed.

4.41 Archaeological Assemblages

Roc de Marsal is a small cave site located in a small tributary valley of the Vézère River in the Dordogne region of southwestern France. Original excavation of the site was carried out by Lafille from 1953 to 1971. The study presented here is based on material from new excavations that took place from 2004 through 2009 (Sandgathe et al. 2011a,b; Turq et al. 2008). A roughly 2m stratigraphic sequence containing 13 stratigraphic layers was recognized, of which Layers 13 through 10 at the base of the sequence represent sterile layers formed through in situ weathering of the limestone bedrock (Sandgathe et al. 2008, 2011b). Thermal luminescence (TL) and optically stimulated luminescence (OSL) dates obtained on sediment samples from these basal layers indicated that initial occupation of the site occurred in Marine Isotope Stage (MIS) 5a (Guérin et al. 2012; Guibert et al. 2009; Sandgathe et al. 2008).

Artifact densities in the Paleolithic layers 9 through 2 are very high, with over 23,000 lithic artifacts greater than 2.5cm in maximum dimension. The lower layers (9-5) contain Mousterian artifact assemblages that are relatively high in Levallois components and include some so-called Asinipodian or small-flake production elements (Bordes 1976; Dibble and McPherron 2006, 2007) and relatively few scrapers. The abundance of

112

fauna remains belonging to forest adapted species, including red deer (*Cervus elaphus*),

roe deer (*Capreolus capreolus*), horse (*Equus sp.*), and wild pig (*Sus scrofa*) (Castel et al.

in Sandgathe et al. 2008) throughout these lower layers indicates a more temperate

climate, although recent OSL and TL dating by Guérin et al. (2012) suggest the

association of these layers with the colder MIS 4.

The upper layers (4-2) saw a change in the lithic assemblage with greater

frequencies of scrapers, including numerous diagnostic Quina scrapers. A dominance of

reindeer (*Rangifer tarandus*) and various vole species, including common vole (*Microtus*

arvalis) and water vole (*Arvicola terrestris*), from these layers indicate a much colder,

drier and more open environment (Marquet in Sandgathe *et al.* 2008). Electron spin

resonance (ESR), TL, and OSL dates from these upper layers suggest correlation with

MIS 4 and 3 (Guérin et al. 2012; Sandgathe et al. 2008).

Pech de l'Azé IV: is one of a complex of four Lower and Middle Paleolithic sites

located in the Dordogne region, about 24km east of Roc de Marsal. The site is a collapsed

cave originally excavated by Bordes (1975) from 1970 to 1977 (McPherron and Dibble

2000). The assemblages examined here come from renewed excavations at the site that

took place from 2000-2003. Eight major Pleistocene layers were identified that in general

matched the sequence identified by Bordes (Turq et al. 2011). The basal layer, Layer 8,

rests directly on bedrock and contains rich Middle Paleolithic materials as well as

numerous superimposed combustion features (Dibble et al. 2009; Goldberg et al. 2012).

The lithic components are marked with high frequencies of scrapers and Levallois

elements. Recent TL dates attributed this basal layer to MIS 5c (Richter et al. 2013). The

overlying Layer 7 represents a solifluction lobe, which indicated by a general lack of

faunal material and a large component of heavily rolled, rounded, or edge-damaged

lithics (Sandgathe et al. 2011b). This layer is capped by a layer of major roof fall, thus

providing further evidence of severely cold conditions during its formation. Layer 6

(subdivided into 6A and 6B) contain lithic elements with high scraper proportions and

noticeable Levallois and Asinipodian components. The faunal record in this layer

indicates a temperate environment with the presence of red deer, roe deer, wild pig, and

beaver (*Castor fiber*), a finding that matches the climatic conditions of MIS 5a (Dibble et

al. 2009; Richter et al. 2013; Sandgathe et al. 2011).

Layer 5 (subdivided into 5A and 5B) saw an increase in reindeer remains while

the presence of roe deer and wild pig decreased or disappeared, signaling the start of a

colder period that likely correlates with the beginning of MIS 4. This interpretation is

supported by a mean date of 68-71 kya from four TL samples (Richter et al. 2013). The

lithic assemblages also contain greater frequencies of scrapers while Levallois elements

declined significantly. In Layer 4 (subdivided into 4A, 4B, and 4C), the trend of change

continues where reindeer becomes the dominant species represented in the faunal record

(Niven 2013). The lithic assemblage, particularly that of Layer 4A, is very rich in

scrapers, including many heavily-reduced forms. The uppermost Layer 3 (subdivided into

3A and 3B) contains an industry that correlates with the Mousterian of Acheulian

Tradition (MTA), with bifaces and backed knives present. Both ESR and AMS dates

from this layer suggest a temporal association with MIS 3 (McPherron et al. 2012; Richter et al. 2013; Sandgathe et al. 2011).

Combe-Capelle Bas: is part of a Paleolithic site complex located in the valley of the Couze River, a tributary of the Dordogne. The site is at the base of a hillside slope below a south-facing limestone cliff approximately 20 km southwest from Roc de Marsal. Early excavation was conducted by Ami from 1926 to 1931. The data presented here come from the excavations of Dibble and Lenoir from 1987 to 1990 (Dibble and Lenoir 1995). In the new excavation, three sectors were established based on Ami's old trench. Sector I is located at the base of the slope and contains a number of stratigraphic layers. Following Roth and Dibble (1998), these layers are grouped into three general units in this study: upper (layers I-IB, I-ICI, I-1C2, I-ID, and I-IDI), lower (layers I-lE, I-2A and I-2B), and I-3. The deposits from Sector II appear to be stratigraphically distinct from those of Sector I and likely shared different depositional histories (Dibble and Lenoir 1995). The six layers from this sector are grouped into two major units – II-3 and II-4 (contains II-4A, II-4B, II-4C, II-4D, and II-4E). Sector III is located at the top of Ami's trench and is represented by several layers. However, only the upper layer (III-1) is included in this study due to the small sample of artifacts within other layers.

The lithic assemblages from all layers at Combe-Capelle Bas share similar typological and technological characteristics. The overall presence of scrapers and notches/denticulates in moderate proportions and the lack of bifaces have led the assemblages to be attributed to the Typical Mousterian industry. However, their

115

technological similarity to Quina Mousterian with an emphasis on thick flake production as defined by Turq (1989, 1992) has been emphasized by Dibble (1995), who further argued that Combe-Capelle Bas assemblages represent a Quina industry with low rates of raw material utilization in terms of scraper production due to the abundance of raw material at the locality. More recent TL dates on samples from Sector I have placed the Combe-Capelle Bas assemblages to MIS 3 (Valladas et al. 2003).

4.42 Computation of the Cortex Ratio

In this study, the calculation of the Cortex Ratio is based on the analysis of all lithic artifacts greater than or equal to 25mm in maximum dimension from the three sites. Only artifacts made on local flint types are considered in order to control for potential variation in original nodule morphology among other raw material types. Among these artifacts, pieces exhibiting alluvial cortex are also excluded to limit the consideration to ones made from locally derived nodules.

The first step is to determine the total cortical surface area in the assemblage, which is only relevant for objects that exhibit cortex. For complete flakes, flake fragments, and retouched pieces, surface area is estimated by multiplying artifact length by width. Length on complete flakes is measured from the point of percussion to the most distal end of the flake, and width is measured at the midpoint of, and perpendicular to, the length axis (Debénath and Dibble 1994). For flake fragments lacking platforms, maximum dimensions are taken instead. For shattered pieces that possess no diagnostic flake features, surface area is calculated in the same way but the value was further

multiplied by two to compensate for the greater artifact surfaces that have the potential to bear cortex (Dibble et al. 2005). For cores and core fragments, an ellipsoid equation is used to account for the three geometric axes of maximum length, width, and thickness (see Douglass et al. 2008; Lin et al. 2010 for detail).

Artifact surface area is then multiplied by the midpoint of the cortex proportion present on each artifact to give an estimate of cortical surface area. In this study, cortex proportion was recorded here with a seven-interval scale: 0%, 1-9%, 10-39%, 40-59%, 60%-89%, 90-99%, 100% (Dibble et al. 2005). Total assemblage cortical surface area is calculated as the sum of the cortex area for each individual artifact. The accuracy of this recording protocol was tested on a set of experimentally produced artifacts (n=77) made from a flint nodule collected close to Roc de Marsal. Cortical surface area on each artifact was digitized and measured using the NextEngine scanner with methodology described elsewhere (Lin et al. 2010). The total cortical surface area estimated from cortex intervals (64,488.29 mm^3) underestimates the scanned area (69,645.43 mm^3) by 7.4%, although the two sets of values are not significantly different (Student's t-test: t=.985, df=76, p=.328). This result differs from that of the study by Lin and colleagues (2010), which showed an 11% overestimation of the scanned value in their cortex calculation. This discrepancy likely resulted from their use of maximum clast dimension for calculating artifact surface area for all artifact types.

The second step is to calculate the expected cortex amount in a given assemblage. This requires an estimate of the number and morphology of nodules that were brought to

the site and reduced to create each of the assemblages. Dibble et al. (2005) suggested the

using the number of cores in an assemblage to estimate the number of nodules (also see

Douglass et al. 2008 and Douglass 2010 for archaeological application, and Douglass and

Holdaway 2011and Douglass 2010 for demonstration of the suitability of this approach

for the Australian case studies). However, because the cores in the study assemblages

here are generally heavily reduced, there was little evidence to suggest a pattern of one

core per worked nodule.

Alternatively, the average size of the originally worked nodules for each of the

three sites can be estimated based on the length of the longest flake present at each site.

This is based on the premise that the largest producible flake is limited by the nodule's

dimensions. Then again, while the maximum length of a nodule constrains the longest

flake length achievable from a nodule, maximum nodule length is likely to be more often

much greater than the length of the longest flake. This trend is likely because, in order for

knappers to fully exploit the longitudinal dimension of a core in a single strike, a tight

combination of platform variables is required (Dibble and Rezek 2009; Lin et al. 2013).

This kind of platform configuration is likely rare at the early stages of reduction when

cortex surface area is at the fullest.

To test the relationship between nodule size and the longest possible flake length,

an experimental assemblage was produced by one of the authors (SCL) using 30 flint

nodules of various sizes (1010-3800g), shapes, and reduction intensity. Prior to reduction,

each nodule's dimensions (maximum length, width, and thickness) and weight were

recorded (Table 1). Nodules were reduced by freehand hardhammer percussion through simple flake removal as well as bifacial and single-surface flaking (*sensu* Sandgathe 2004; but see Eren and Bradley 2009). A total of 2,502 flakes and flake fragments larger than 25mm were produced. The length of the longest flake from each nodule reduction set was used as the dimensional proxy to estimate original nodule volume (Table 1). Linear regression models showed best results when the longest flake length was used as proxy for nodule width, which produced an average 17% underestimation of actual volume (s.d.=27%; regression: standardized coefficient (β) =.43, p<.05). This method outperformed models based on nodule length (underestimation by 42.1% with s.d.=19%; regression: β=.47, p<.05) and thickness (overestimates by 42.34% with s.d.=54%; regression: β=.84, p<.001). It is true that, in some cases, flake lengths would mimic nodule length more than width, such as blade production, or when nodules are of cylindrical or conical shape. However, it is argued here that this general correlation between longest flake length and original nodule width is largely related to solid geometry and likely to hold for most stone morphologies and flake-based production techniques, including those of the Mousterian assemblages studied here. Therefore, the length of the longest flake from each of the three study assemblages was taken here as an estimate of original nodule width for the purpose of approximating nodule volume.

Nodule volume (cm³)	Max. length (mm)	Max. width (mm)	Max. thick. (mm)	Length of longest flake (mm)	Nodule volume (cm³)	Max. length (mm)	Max. width (mm)	Max. thick. (mm)	Length of longest flake (mm)
1437.8	239.0	110.6	99.3	97.6	811.2	150.3	137.1	79.5	109.5
433.5	194.6	85.9	43.4	69.7	974.2	195.1	119.3	67.5	87.8
888.4	170.5	136.0	72.0	82.7	987.1	174.6	131.1	80.5	125.9
815.5	191.4	120.3	71.5	103.6	699.6	151.5	151.5	51.6	57.6
845.5	178.0	97.1	87.4	56.3	974.2	181.6	138.9	70.2	76.5
613.7	155.0	105.1	65.0	91.4	961.4	193.8	128.0	87.2	110.1
1111.6	170.5	144.0	79.1	112.6	712.4	221.2	135.3	71.1	91.2
639.5	151.4	111.5	72.8	76.3	1257.5	223.2	114.7	87.0	103.7
918.5	188.9	134.5	65.3	90.4	905.6	154.5	96.5	93.0	106.0
802.6	170.8	130.6	77.6	107.6	884.1	154.6	133.4	79.3	104.6
506.4	211.9	105.7	39.7	94.4	622.3	155.0	107.9	69.1	67.9
553.6	207.8	118.3	49.2	127.1	1630.9	224.6	147.3	90.1	98.2
1047.2	201.9	141.1	90.3	99.5	1326.2	236.1	115.4	105.3	90.8
763.9	193.6	128.9	61.2	74.8	446.4	170.9	85.4	40.3	72.4
1304.7	214.6	111.9	90.0	105.9	712.4	178.5	113.2	51.8	64.4

Table 1 – Basic descriptions of the thirty experimentally reduced nodules used to test the use of longest flake length to approximate original nodule volume.

To determine nodule size, the derived axial estimates of nodule width needs to be related to the overall nodule volume. Fortunately, flint occurs naturally in close proximity to all three sites, and the majority of artifacts from each assemblage are attributed to these locally derived flint types (Turq et al. 2008, 2011). This finding matches the general observation that raw materials represented in Middle Paleolithic sites of Western Europe tend to be dominated by local rock types, most coming from within a radius of 4-5km from the site (Feblot-Augustins 1999; Fernandes et al. 2008; Geneste 1985; Kuhn 1991; Turq 1992, 2000; Turq et al. 2013). At Roc de Marsal, local flint is found on the adjacent limestone plateau. For Pech de l'Azé IV, Senonian flint nodules exist along the immediate hillside slope and nearby valley floor (Turq et al. 2011). At Combe-Capelle

Bas, Campanian-Senonian flint outcrops are found along the valley slope where the site is located (Turq 1995). Samples of these local flint types were collected to generate regression models for estimating nodule size based on width (Figure 1). Because flint nodules are mostly buried at these localities and difficult to find, the survey could not be conducted systematically for the goal of maximizing sample size. Rather, emphasis was placed on obtaining a representative size range of the naturally occurring nodules at each locality.

A total of 26 nodules (175-1505g) were collected from Roc de Marsal and 23 (197-1443g) from Pech de l'Azé IV. At Combe-Capelle Bas, samples were taken from one subsurface outcrop exposed by a recent road cut near the site (n=14; 539-7579g). However, because these flints were freshly broken, only nodules with more than 60% cortex were used as they retain greater resemblance to those that occur naturally; though this limits the sample size considerably (n=6; 539-2157g). These cortical nodules are mostly of oblong tubular form (Figure 1). All raw material samples were digitized into three dimensional models by the same procedure mentioned above, and volume, surface area, and cortical surface area were computed. A comparison between nodule mass and scanned volume provided an average density constant of 2.34, which was then used to convert artifact mass to volume.
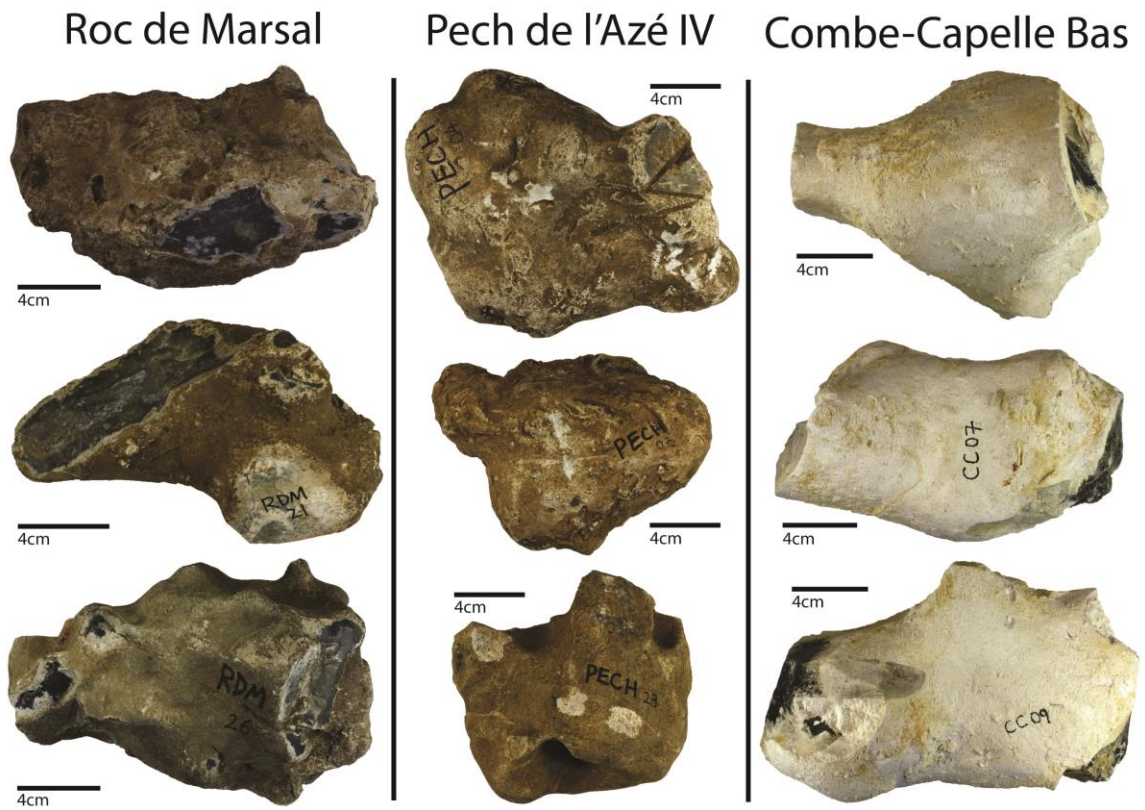
Figure 1 – Examples of flint nodules collected nearby the three sites.

Because the axial dimensions of these flint nodules were found to strongly correlate with nodule volume, a set of linear regressions were generated for predicting original nodule volume based on maximum nodule width for Roc de Marsal and Pech de l'Azé IV:

$$\textit{Roc de Marsal: } V=(.573 \times W + 10.944)^3 \ (n=26, \beta=.84, p<.001)$$

$$\textit{Pech de l'Azé IV: } V=(.495 \times W + 19.113)^3 \ (n=23, \beta=.88, p<.001)$$

where V is the reconstructed nodule volume and W is maximum nodule width. The length of the longest complete flake from the site was then input into the respective regression model as an estimate of nodule width to calculate original nodule volume. While it is true that nodule size may have varied through time between layers, this measure allows an estimate of the largest nodule that could have been utilized at each site.

For Combe-Capelle Bas, a different approach for estimating nodule volume was employed since, in these assemblages, there are a number of cores that are much larger (in maximum length) than the longest flake. This likely occurs because the site is situated on or immediately adjacent to the outcrop. Therefore, the largest core was taken instead for establishing the approximation. The following regression model was derived from the nodule samples collected at Combe-Capelle for nodule volume based on nodule length:

$$V=(.599 \times L - 10.23)^3 \ (n=6, \beta=.84, p<.05)$$

where V is, again, the reconstructed nodule volume and L is maximum nodule length. The longest core length from the assemblage was input as nodule length in order to calculate the largest nodule possible for the site. The number of nodules used for constructing this regression is small (n=6). Therefore, it is expected that the estimated nodule volume for Combe-Capelle Bas will have greater error ranges than the estimates of the other two sites. The reconstructed nodule volume for each site was then used to divide the total artifact volume of the respective assemblages to arrive at the estimated number of nodules.

The third step in the Cortex Ratio calculation is determining the expected amount of cortex per nodule. This is done by inputting the reconstructed nodule volume into appropriate equations that approximate the surface area of a geometric solid based on a given volume. A comparison between scanned cortex surface area of the collected nodule samples and those established from different geometric solid models, including sphere, cylinder, and cube (Dibble et al. 2005), showed that the surface area to volume relationship of a cylinder [Surface Area=4π(Volume/π)$^{2/3}$] best summarizes the average characteristic of Pech de l'Azé nodules, with an average underestimation of 0.29% (s.d.=6.6%; Pearson correlation: n=23, r=.99, p<.001). On the other hand, the cube equation (Surface Area=$6\times$Volume$^{2/3}$) was best for nodules collected from Roc de Marsal, with an underestimation of 5.82% (s.d.=6.41%; Pearson correlation: n=32, r=.99, p<.001), and Combe-Capelle Bas, with an underestimation of 0.51% (s.d.=4.34%; Pearson correlation: n=5, r=.97, p<.05). It should be noted that the use of these models does not imply that the shapes of raw materials resemble these standardized geometric

124

solids, although in some cases this may be true (Douglass 2010; Douglass et al. 2008). In this study, it simply means that the surface area to volume relationships of these solids approximates those of natural stones.

The reconstructed nodule surface area represents the amount of cortex on a stone if cortex coverage is 100%. However, unlike alluvial stones that tend to be completely cortical, the cortical surfaces of flint nodules derived from local limestone formations tend to contain varying degrees of old, non-cortical surfaces that are unrelated to anthropogenic activities but that can be confused with newly exposed (rather than cortical) surfaces when quantifying the cortex on the artifacts of an assemblage (see Figure 1). An assumption that the flint nodules reduced at a site all began in complete cortical forms would therefore overestimate the expected amount of cortex that should be present in the assemblage. To correct for this assumption, the estimated cortex is corrected by subtracting the naturally occurring non-cortical surfaces on these nodules.

From our flint sample, nodules from Roc de Marsal on average have 18.5% (s.d.=13.7%) non-cortical surfaces and Pech de l'Azé nodules have 9.2% (s.d.=8.6%). Since nodules collected from Combe-Capelle were freshly broken, this consideration was made on the possible non-cortical surfaces that may have existed when flint in tubular outcrop form were broken up in the past. Tubular pieces in our sample exhibit on average 30.6% (s.d.=9.9%) non-cortical surface. These average non-cortical proportions were subtracted from the estimates of nodule surface area at each site to arrive at the expected cortical surface area per nodule.

125

Total expected cortex for an assemblage was computed by further multiplying this cortex estimate by the estimated number of original nodules as described above. Dividing the total observed cortex of a given assemblage by the total expected cortex results in the Cortex Ratio. A 75% confidence interval was also determined for each Cortex Ratio to account for the error range in nodule volume estimation. This level was chosen because higher confidence intervals would return negative volume estimates for the Combe-Capelle Bas assemblages, which makes calculation of ratio values impossible.

4.5 The Experimental Dataset and Archaeological Resampling

Two approaches were employed to examine the null hypothesis that a given archaeological Cortex Ratio is statistically equivalent to a ratio from an assemblage that is "complete" (i.e., where cortical nodules were brought in and reduced and where no artifacts were subsequently removed or added), which in theory should have a ratio of 1. For the first approach, the hypothesis tested is whether a given archaeological Cortex Ratio is different from a ratio of 1. To do this, each archaeological assemblage was bootstrapped to generate a sampling distribution that provides an estimate of the shape and range of Cortex Ratio distribution. Sampling was done *with* replacement, meaning that after an artifact was randomly selected, it was placed back into the sampling population and had an equal chance of being selected again. Bootstrapping was conducted in R software package (R Core Team 2013) and contains three steps: (a) randomly resample from the archaeological assemblage with replacement a sample of artifacts of size equal to the original assemblage; (b) calculate the Cortex Ratio on this sample; and (c) repeat these steps 10,000 times. Ten thousand iterations was chosen in

order to have a maximum detectable significance of p=.0001 and to capture the potentially high variance of the Cortex Ratio due to the arithmetic nature of ratios where the numerator and denominator affect the value unequally (discussed below).

Because the distribution is nonsymmetrical, it is necessary to look at one side of the curve or the other depending on whether the archaeological Cortex Ratio is less than or greater than 1. If the archaeological Cortex Ratio is below 1, then the one-tailed probability for the archaeological Cortex Ratio to equal 1 is represented by the relative occurrence of ratios in the sampling distribution that are equal to or greater than 1 (Figure 2). On the other hand, if the archaeological Cortex Ratio is greater than 1, then the one-tailed probability is represented by the relative frequency of cases equal to or smaller than 1. Because the hypothesis is non-directional (i.e., the direction of difference *not* predicted), the two-tailed probabilities were estimated by multiplying the one-tailed probabilities by two. An alpha level of 0.05 was employed for this test as well as all other tests described below.

Archaeological Cortex Ratio = .75

n ratios with values equal to or greater than 1

mean = .75   1

One-tailed probability = $n/10000$

Two-tailed probability = $(n/10000) \times 2$
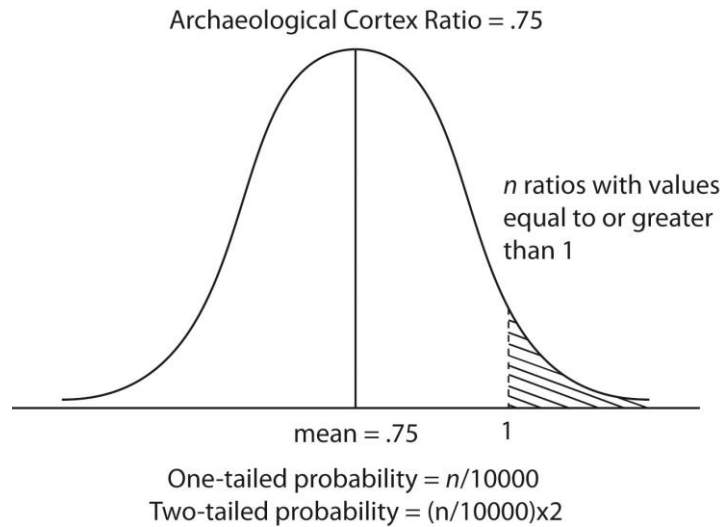
Figure 2 – A hypothetical bootstrap sampling distribution of an archaeological assemblage with a Cortex Ratio of .75. Because the archaeological Cortex Ratio is below one, the one-tailed probability is determined by the occurrence of resampled ratio values equal to or greater than 1 (represented by the shaded area).

This first approach provides a relatively efficient way of evaluating the probability that a given archaeological assemblage can produce a Cortex Ratio of 1. However, this method relies solely on the archaeological sample itself and is unable to account for the potential range of variation in the Cortex Ratios producible from a "complete" assemblage. In other words, although an assemblage may have unity between cortex and volume, its Cortex Ratio may not be exactly 1 due to sampling error and other factors.

To take this issue into consideration in the process of establishing statistical significance, we employed a second approach using a large collection of experimental data containing multiple reduction sets to generate sampling distributions of Cortex Ratios at different sample size levels. These sampled Cortex Ratios represent the range of

values obtainable from assemblages that retain all products of nodule reduction. The data include the previously described reduction sets made on 30 flint nodules, and also two other experimental sets from previous studies of Dibble et al. (2005) and Douglass (2010). The set produced by Dibble and colleagues (2005) was made on 33 nodules of chert and obsidian with varying forms and sizes (350-4350g). Nodules were reduced by multiple knappers using both hard and soft hammer through a variety of technologies, although most were simple flake removals without significant core preparation. The Douglass (2010) set was made on 29 silcrete nodules (107-1535g), mostly spherical in shape, with freehand hardhammer percussion and also followed a general flake production technique without formal core preparation techniques. When combined, the three experimental datasets included 236 nodule reductions and a total of 9,524 artifacts with maximum length equal to or over 25mm. Reduction intensities of the nodules vary from 2 to 125 flakes (with platform) per nodule.

It may appear intuitive to argue that the experimental dataset should be produced primarily through reduction strategies that are characteristic of the Middle Paleolithic assemblages studied here, namely Levallois and prepared core techniques. However, because these assemblages cover a range of dissimilar technological characteristics (Levallois, Quina, MTA), it is difficult to determine in what way the experimental data should be generated to best resemble the archaeological reality. At a deeper level, the issue is whether the experimental data are comparable to the archaeological data, namely, if the two share similar structural variables that are relevant to the metric calculated, i.e., the Cortex Ratio. Cortex Ratio is largely a function of the geometric composition of the

assemblage, expressed in the forms of flake size and cortex distribution. However, if one wants to make the geometric composition of an experimental dataset to match that of an archaeological assemblage, then the resulting experimental Cortex Ratio would inevitably be constrained by the need to resemble that of the latter.

The point here is that the purpose of the experimental data is not to replicate the archaeological assemblage, but to serve as a reference for null hypothesis testing. We recognize that such a procedure requires the assumption that the experimental dataset captures a range of variability akin to those exhibited in the archaeological assemblages studied here. The reason for compiling such a large dataset with diverse nodule morphology, reduction intensity, and knapping technique is thus to attain the largest possible range of variation in reduction sets. The difference between some of the raw materials used in the experimental dataset and that of the archaeological assemblages in question here would have little impact on the utility of the experimental data because the way that stone volume and cortex are distributed among flaked products is similar across raw material types. The main feature of the experimental data that is of interest here is that it contains all of the products of nodule reduction; therefore, in theory, it should have a Cortex Ratio of 1. In fact, a Cortex Ratio of 1.00 was calculated from the combined experimental dataset by using the known average original nodule weight and a spherical model to approximate nodule shape (Douglass et al. 2008; Lin et al. 2010).

Treating the compiled experimental data as a population with known geometric parameters, a sampling distribution of Cortex Ratios was generated by a Monte Carlo

sampling approach that randomly draws samples from the experimental population over a large number of instances. Sampling was again done *with* replacement. Consequently, the samples of artifacts drawn share a distinct structure from the experimental data but attain a mean Cortex Ratio that approximates one. The test hypothesis, which again is non-directional, is whether or not the archaeological cortex ratio is different than the mean ratio of the experimental dataset and based on the shape of the sampling distribution formed by the repeated iterations.

The sampling process was carried out by a similar R routine as described before and contains the following three steps: (a) randomly sample with replacement a group of artifacts from the experimental data with a size equal to the archaeological assemblage in question; (b) calculate the Cortex Ratio of the sample; and (c) repeat this routine 10,000 times. This routine creates a sampling distribution of Cortex Ratios of a given sample size corresponding to the archaeological assemblage. For archaeological Cortex Ratios lower than 1, the one-tailed probability is represented by the relative occurrence of ratios in the sampling distribution that are equal to or greater than the archaeological ratio. On the other hand, if the archaeological Cortex Ratio is greater than 1, then the one-tailed probability is represented by the relative frequency of cases equal to or smaller than the archaeological ratio. Similar to the bootstrapping test above, two-tailed probabilities were estimated here by multiplying the one-tailed probabilities by two.

The second research issue of this study is to compare two archaeological Cortex Ratios with statistical confidence. This can be translated to a test of the null hypothesis

that the two archaeological samples were drawn from the same population. This is achieved with a separate R routine that performs a permutation test, which is a randomization test that creates the null hypothesis distribution of a test statistic by permuting the combined data of the two compared groups. However, since there may be too many possible orderings to the combined data, an exact test is adopted here instead of Monte Carlo sampling to generate reference distributions over limited number of permutations.

Here, the method creates an empirical sampling distribution of the differences between the two archaeological Cortex Ratios in question. This routine consists of the following steps – a) calculate the difference between the Cortex Ratios of two archaeological assemblages, b) combine the two assemblages, c) randomly draw a sample *without* replacement of artifacts of the same size as one of the assemblages in question, d) assign the rest of the artifacts (the previously unselected ones) in the pooled data to represent the other assemblage in question, e) compute the difference between the Cortex Ratios of the two samples of artifacts, f) repeat steps c through e 10,000 times. The Cortex Ratio is calculated with the procedure used for the archaeological study assemblages as described above. Because the difference between the two archaeological Cortex Ratios in question could be either negative or positive depending on which way the subtraction goes (e.g., the difference between ratios of 0.5 and 0.8 can be either 0.3 or -0.3), the probability that the two archaeological Cortex Ratios were drawn from the same population of artifacts is based on the relative frequency of resampled Cortex Ratio

differences both greater than or equal to the positive difference and less than or equal to

the negative difference (Figure 3).

Archaeological Cortex Ratios 0.5 vs. 0.8
Difference = ±0.3

n1 ratios with values
equal to or less
than -0.3

n2 ratios with values
equal to or greater
than 0.3

-0.3

mean = 0

0.3

p=(n1+n2)/10000
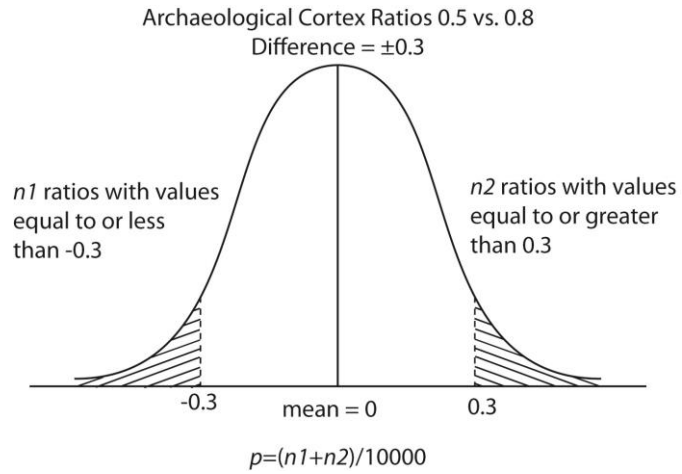
Figure 3 - A hypothetical sampling distribution of the differences between two archaeological assemblages with Cortex Ratios of 0.5 and 0.8 generated by the permutation test. The shaded areas represent the occurrence of resampled Cortex Ratio differences that are more extreme than the actual difference between the two archaeological ratio values.

4.6 Results

Table 2 shows the Cortex Ratios and other associated variables of the three sites. The probabilities for Cortex Ratios to equal 1 or to be drawn from a complete assemblage differ considerably between the bootstrap and Monte Carlo approaches. According to the bootstrap approach, over half of the layers examined here have Cortex Ratios that are statistically different from a ratio of 1; however, results from Monte Carlo sampling suggest they fall within the 95% confidence interval. This has likely occured because the experimental dataset produces more variability than a single archaeological assemblage. As a result, although these layers did not produce a ratio of 1 in the bootstrap distribution, their respective ratios in fact fall within the range of values that can be produced from complete assemblages with chances over the level of statistical significance. Because of this, the Monte Carlo results provide a more conservative perspective on the testing of the null hypothesis. Considering these issues, only the results from the Monte Carlo test are used in the following analyses.

Layer	Total Artifact Count	No. of Expected Nodules	Observed Cortex Surface (mm^2)	Expected Cortex Surface (mm^2)	Cortex Ratio (± 75%)	Bootstrap p (equal to ratio of 1)*	Monte Carlo p (from assemblage with ratio of 1)*
Roc de Marsal (Length of longest flake=116.5mm; Predicted Nodule Mass=1097.4g)							
2	685	12.8	268158	376339	.71 ±.09	**<.0001**	**.019**
3	174	3.1	74618	92255	.81 ±.11	**.004**	.43
4	2485	32.4	650745	956136	.68 ±.09	**<.0001**	**<.0001**
5	1780	17.4	426358	512845	.83 ±.11	**<.0001**	**.042**
6	501	5.3	145098	154886	.94 ±.15	.087	.70
7	2758	25.2	680740	742792	.92 ±.12	**.0001**	.28
8	2410	25.7	734792	758438	.97 ±.13	.40	.71
9	4651	40	1198649	1177759	1.02 ±.14	.40	.75
Pech de l'Azé IV (Length of longest flake=129.5mm; Predicted Nodule Mass=1347.5g)							
3A	1547	9.6	454295	354296	1.28 ±.14	**<.0001**	**.042**
3B	2762	14.8	650882	545835	1.19 ±.13	**<.0001**	**.041**
4A	216	2	62985	72222	.87 ±.09	.17	.57
4B+C	647	5.6	216923	207359	1.05 ±.12	.68	.80
5A	1366	8.1	371261	299270	1.25 ±.13	**<.0001**	.076
5B	622	5.2	193922	193424	1.00 ±.11	.96	.98
6A	2309	17.8	660788	656330	1.01 ±.11	.82	.92
6B	2304	14.2	576487	522843	1.10 ±.12	**.0002**	.31
7	2784	18.7	572112	690488	.83 ±.09	**<.0001**	**.015**
8	2076	9	336748	332025	1.01 ±.1	.72	.95
Combe-Capelle Bas (Length of longest core=246mm; Predicted Nodule Mass=6033.3g)							
I-upper	2167	11.5	589292	954442	.60 ±.45	**<.0001**	**<.0001**
I-lower	2222	31.7	2338540	2630714	.86 ±.66	**.0002**	.071
I-3	109	0.9	61774	78710	.76 ±.58	.29	.43
II-3	207	0.5	29618	38296	.75 ±.57	.087	.27
II-4	1484	5.9	366107	487072	.73 ±.56	**<.0001**	**.0012**
III-1	525	1.9	74216	155806	.46 ±.35	**<.0001**	**<.0001**

* Two-tailed probability

Table 2 – Summary of Cortex Ratios and other results for the study assemblage layers. Bootstrap p represents the probability derived from bootstrap distribution of the archaeological sample itself. Monte Carlo p denotes the probability derived from comparing the archaeological Cortex Ratio to the test distribution generated from the experimental dataset. Bold values represent probabilities that are statistically significant.

Most layers at Pech de l'Azé IV and the lower layers of Roc de Marsal have ratio

values close to 1. Conversely, the two upper layers of Pech de l'Azé IV show an excess

of cortex while those at Roc de Marsal indicate a general deficit in cortex. Combe-

Capelle Bas shows a general pattern of underrepresenting cortex throughout the

sequence; though three of the six layers did not exhibit ratios that are statistical different

from 1. The probability for a given Cortex Ratio to be not statistically different from a

complete assemblage is partly affected by assemblage size. This outcome is particularly

clear with Layer 3 of Roc de Marsal (n=174), Layer 4A of Pech de l'Azé IV (n=216) and

Layer I-3 and II-3 from Combe-Capelle Bas (n=109 and 207 respectively), where a below

1 Cortex Ratio can be achieved through sampling with considerably higher probability

due to the small sample size of these layers.

That being said, there are large assemblages, such as many of the lower layers at

Roc de Marsal and Pech de l'Azé IV that are not statistically different from complete

assemblages. This suggests these assemblages have not been altered in ways that changed

their initial cortex composition. It should also be noted that the error bars associated with

Combe-Capelle Bas Cortex Ratios are considerably larger than the other two sites due to

the greater uncertainty in the original nodule morphology estimate.

Figure 4 shows examples of the generated Cortex Ratio sampling distributions

and their comparison with archaeological ratio values. The sampling distributions are

slightly asymmetrically skewed towards above 1 values. This is likely a result of the

arithmetic nature of the ratio measure, where the range of variability for above and below

1 ratios is unequal. Specifically, while below 1 values can only range between zero and one, there is no upper bound for above 1 ratios. That said, the sampling distribution and confidence interval does narrow down and become more symmetrical as sample size increases.
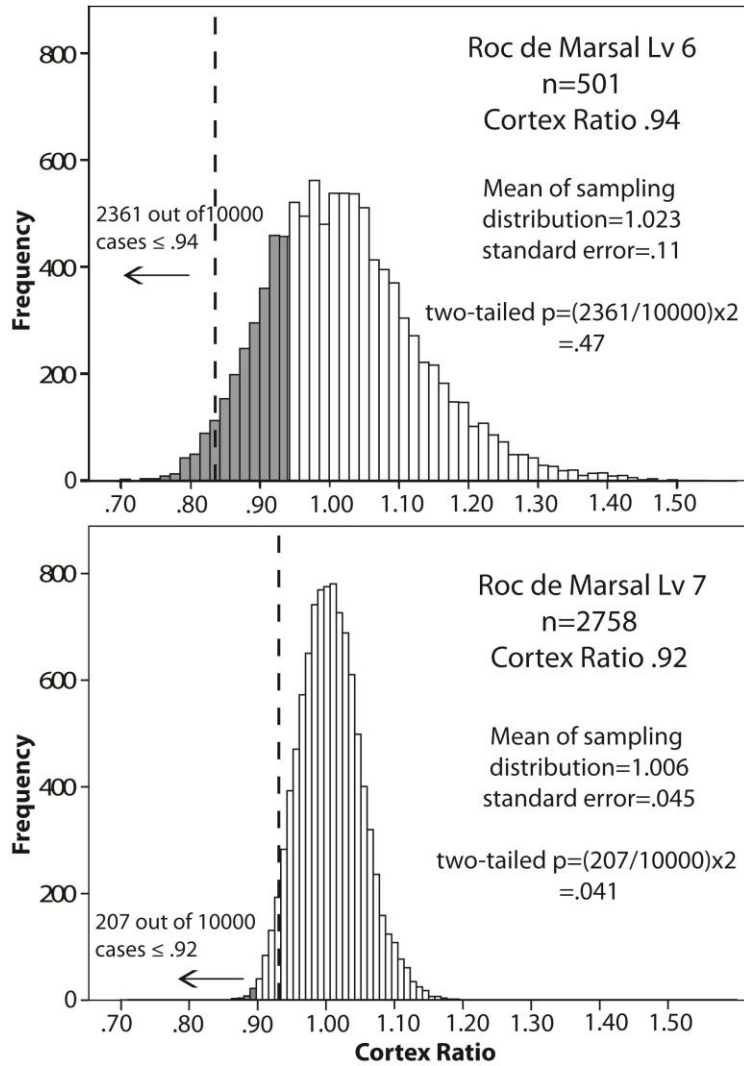
Figure 4 – Sampling distributions derived from the experimental data with sample sizes equivalent to those of Roc de Marsal Layers 6 and 7. Dashed lines indicate the 95% confidence intervals of the distributions. The grey areas represent sampled Cortex Ratios equal to the archaeological ratio or of more extreme values.

Tables 3-5 show the results of the permutation test at the three sites. Figure 5

illustrates examples of sampling distributions derived from permutation test and their

comparison to archaeological Cortex Ratio differences. Again, the issue with sample size

is clear, as layers with small assemblage size, such as Layer 3 at Roc de Marsal, do not

show statistical significance with any other layers due to the wide variance in its

sampling distribution. For Roc de Marsal, the ratios for Layers 9-6 at the bottom half of

the stratigraphic sequence are relatively similar and close to one. However, as we move

up the sequence, Layers 5, 4, and 2 all exhibit ratio values that are significantly lower

than the layers below. If we exclude Layer 3 because of small sample size, then the

Cortex Ratios of each layer do not share significant differences with adjacent layers,

except for Layer 4 which has a ratio that is significantly lower than all others. This

pattern indicates a gradual decline in assemblage cortex over time and the tendency for

cortex underrepresentation in the layers higher up in the stratigraphy.

Layer	2 (.71)	3 (.81)	4 (.68)	5 (.83)	6 (.94)	7 (.92)	8 (.97)
2 (.71)	-	-	-	-	-	-	-
3 (.81)	.29	-	-	-	-	-	-
4 (.68)	.50	.12	-	-	-	-	-
5 (.83)	.032	.81	**<.0001**	-	-	-	-
6 (.94)	**.0047**	.31	**<.0001**	.24	-	-	-
7 (.92)	**.0003**	.29	**<.0001**	**.032**	.87	-	-
8 (.97)	**.0001**	.16	**<.0001**	**.0043**	.45	.27	-
9 (1.02)	**<.0001**	**.049**	**<.0001**	**<.0001**	.089	**.0021**	.25

Table 3 – Permutation test results indicating the probabilities for corresponding layers to be from the same assemblage at Roc de Marsal. Bold values represent statistically significant layers.

Layer	3A (1.28)	3B (1.19)	4A (.87)	4B+C (1.05)	5A (1.25)	5B (1.00)	6A (1.01)	6B (1.1)	7 (.83)
3A (1.28)	-	-	-	-	-	-	-	-	-
3B (1.19)	.068	-	-	-	-	-	-	-	-
4A (.87)	**.0003**	**.0026**	-	-	-	-	-	-	-
4B+C (1.05)	**.0027**	**.044**	.15	-	-	-	-	-	-
5A (1.25)	.64	.3	**.0059**	**.025**	-	-	-	-	-
5B (1.00)	**<.0001**	**.0062**	.21	.59	**.0028**	-	-	-	-
6A (1.01)	**<.0001**	**<.0001**	.2	.57	**<.0001**	.95	-	-	-
6B (1.1)	**<.0001**	**.03**	**.023**	.42	**.0073**	.11	**.03**	-	-
7 (.83)	**<.0001**	**<.0001**	.62	**.0002**	**<.0001**	**.002**	**<.0001**	**<.0001**	-
8 (1.01)	**<.0001**	**.0001**	.23	.68	**<.0001**	.87	.88	.059	**<.0001**

Table 4 – Permutation test results indicating the probabilities for corresponding layers to be from the same assemblage at Pech de l'Azé IV. Bold values represent statistically significant layers.

Layer	I-upper (.60)	I-lower (.86)	I-3 (.76)	II-3 (.75)	II-4 (.73)
I-upper (.60)	-	-	-	-	-
I-lower (.86)	**<.0001**	-	-	-	-
I-3 (.76)	.31	.53	-	-	-
II-3 (.75)	.20	.38	.97	-	-
II-4 (.73)	**.021**	**.029**	.87	.87	-
III-1 (.46)	.074	**<.0001**	.11	**.022**	**.0021**

Table 5 – Permutation test results indicating the probabilities for corresponding layers to be from the same assemblage at Combe-Capelle Bas. Bold values represent statistically significant levels.
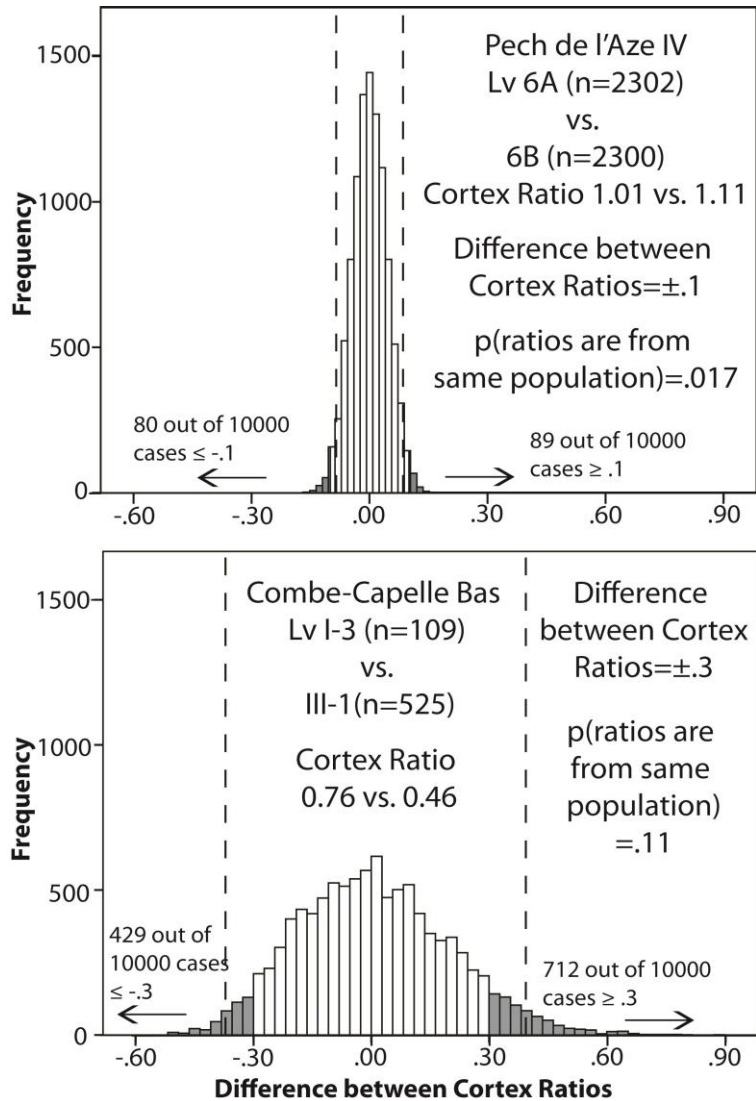
Figure 5 – Sampling distributions derived from permutation tests for comparing Cortex Ratios between Pech de l'Azé IV Layers 6A and 6B, and Combe-Capelle Bas Layers I-3 and III-3. The relative frequency of values occurring at the same level as the Cortex Ratio difference or of more extreme values provide the probability for such difference to be significant. In the two cases shown here, Layers 6A and 6B at Pech de l'Azé IV have Cortex Ratios that are statistically different despite having a difference of just .1. In contrast, the Cortex Ratios for Layer I-3 and III-3 at Combe-Capelle Bas differ markedly in value but share no statistical difference.

140

The permutation results of Pech de l'Azé IV are less straight forward. Several layers have ratios that are relatively similar in value (Layers 4B+C, 5B, 6A, 6B, and 8), and the permutation tests indicate that most of them are not statistically different, with the exception of the difference between 6A and 6B. Unlike Roc de Marsal, most layers here exhibit significant differences in the Cortex Ratio of adjacent layers. Layers 3A, 3B, 5A in the upper half of the sequence stand out for having above 1 ratios that are statistically identical from each other; though the ratio of 5A has been shown to be not statistically different from 1. This pattern suggests that the top two layers of Pech de l'Azé IV contain greater amounts of cortex relative to assemblage volume, while, with the exception of Layer 7, the rest of the layers are statistically indistinguishable from assemblages that contain unity between cortex and volume. However, in comparison to Roc de Marsal, the pattern at Pech de l'Azé IV appears to be more varied through time, being marked with greater fluctuations between layers instead of a more gradual trend.

At Combe-Capelle Bas, Sector III has a Cortex Ratio that is markedly lower than all other layers while the two layers in Sector II share similar ratios. In Sector I, the Cortex Ratios for the three layers are more varied and I-upper and I-lower differ significantly.

Figure 6 graphically presents the Cortex Ratio results of the three sites. Ratios for Roc de Marsal and Pech de l'Azé IV are ordered in stratigraphic sequence with correspondence to fauna assemblage composition. This is done instead of using MIS stages because of the uncertainties in the current dating of the two sites. At Roc de

141

Marsal, layers with ratio values close or equal to 1 have faunal assemblages dominated by red deer, indicating an overall temperate and forested landscape during these time periods. Upper layers with below 1 ratios, on the other hand, are dominated with reindeer remains, and likely correspond to cold and open environments. Pech de l'Azé IV layers at the lower half of the sequence are dominated by red deer remains and also exhibit Cortex Ratios approximating 1. This is especially true if we exclude Layer 7 due to its having suffered from solifluction (Goldberg et al. 2012; Sandgathe et al. 2011b), which may have influenced assemblage cortex composition in some way by removing and/or modifying artifacts of certain shape or clast size. The stratigraphically higher layers have faunal assemblages with higher proportions of reindeer remains. However, only the two upper most layers exhibit Cortex Ratios that are statistically different from 1. Combe-Capelle Bas has been dated to MIS 3, but it is less clear how the three separate excavated sectors relate to each other in terms of chronology due to their distinct depositional histories (Dibble and Lenoir 1995). Therefore, it is difficult to organize Cortex Ratios at Combe-Capelle Bas into a unified chronological framework other than considering them together for the period of MIS 3.
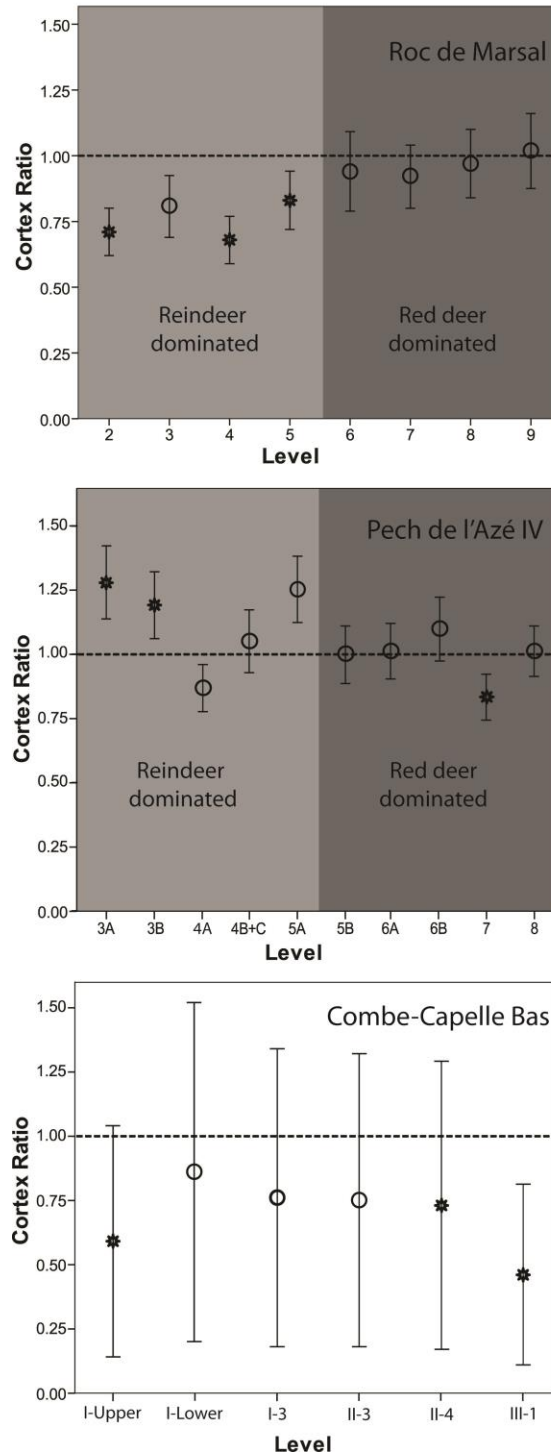
Figure 6 – Cortex Ratios by layer at Roc de Marsal, Pech de l'Azé IV and Combe-Capelle Bas. Circles represent ratio values not statistically different from 1 while stars represent those that differ significantly from 1. The dotted line is reference for a Cortex Ratio of 1. Error bars indicate 75% confidence interval from nodule shape estimation.

4.7 Discussion

The cortex approach differs from traditional methods of measuring artifact transport, such as raw material sourcing and artifact design theory, which largely rely on the presence of specific artifact forms or traits. For example, sourcing requires the occurrence of stone types attributable to localized primary or secondary deposits (Fernandes et al. 2008; Hughes 1998; Jia et al. 2010; Shackley 1998; Sheppard et al. 2011; Montet-White and Holen 1991 and papers therein). Likewise, models of artifact design and technological organization depend on the presence of specific artifact types, including formal or retouched forms (Andrefsky 1994; Bamforth 1986; Kelly and Todd 1988; Nelson 1991). In many instances, however, the majority of artifacts within an archaeological assemblage was not made on stones of demonstrably non-local origin and tends to lack significant retouch. Conjoined artifacts from different sites are another possible means for demonstrating movement between sites, but they also tend to be rare and their identification is heavily constrained by time and expertise (Close 2000). Therefore, the representativeness of these observations in relation to the overall pattern of artifact transport at an assemblage scale remains open to question.

More specifically, while the existence of non-local material or certain artifact types may indeed indicate forms of transport, it is less certain whether the absence of these material markers signal the opposite pattern of no artifact movement. This issue illustrates the danger of false negatives in these conventional approaches which potentially have significant ramifications for the interpretation of past mobility and technological organization. In contrast, the Cortex Ratio quantifies the overall

144

composition of the assemblage regardless of artifact type or form. Because the measurement is a ratio variable standardized by assemblage size and operates solely on the principles of solid geometry, the applicability of the method is not constrained by variability introduced through technological differences, such as knapping technique and reduction intensity.

A critical aspect of the calculation of the Cortex Ratio is the cortex-to-volume proportion estimated for the original knapped nodules, which serves as the main reference for the quantification of assemblage cortex. While Dibble et al. (2005) proposed the use of geometric solid models for such approximation, other ways for calculating the relationship can nevertheless be applied. The flexibility in the way that this estimate can be achieved means the cortex quantification approach can be applied to a wide array of archaeological settings. On the other hand, because this estimate is contingent on the geometrical qualities of local raw material, it would be inappropriate to expect a methodological procedure developed under one circumstance to work in another. Instead, it is important to validate and test appropriate approximations of nodule geometry through contextual observation and experiments. The study here employed the length of the longest flake present in an assemblage as proxy for the axial dimension of the largest nodule that was used at the site. Geometric models of nodules were then generated from raw material samples obtained through survey. Given that actual nodule sizes in the past were likely larger than what is estimated here, the values provide a conservative baseline that is empirically supportable, although they probably underestimate the expected amount of cortex by some extent.

Once the Cortex Ratio of an assemblage is derived, its interpretation faces two main challenges. First, it is difficult to discern at face value whether an archaeological ratio is statistically different to those generated from complete assemblages that contain all products of nodule reduction. The ability to establish statistical significance between a given Cortex Ratio and the ratios of complete assemblages therefore provides a vital mean for interpreting assemblage cortex composition. As shown earlier, the probability for a given Cortex Ratio to be derived from a complete assemblage is largely dependent on sample size. Here, a large experimental dataset consisting of complete reduction sequences was employed to create sampling distributions of Cortex Ratios based on different sample sizes. This, in turn, allowed us to determine the probability for an archaeological assemblage with a given size and ratio value to be derived from a population that is at unity between artifact cortex and volume.

This approach, however, requires repeated sampling of Cortex Ratios from the experimental data for any given archaeological assemblage. A possible alternative method is to compute the confidence interval of the sampling distribution based on the size of the archaeological assemblage under question. Figure 7 shows the plotted mean and standard error of sampled Cortex Ratios from the experimental dataset with varying log-transformed sample sizes. The pattern where small samples have much higher Cortex Ratio may be explained by the different rate of increase between surface area and volume for geometric solids. Specifically, surface area increases as a square while volume increases as a cube. As a result, when an assemblage has only a few artifacts, there is likely to be proportionally more surface area than artifact volume present. However, the

146

plot does show that the Cortex Ratio mean becomes increasingly stable after sample size

reaches ~500 (log-transformed=6.2).
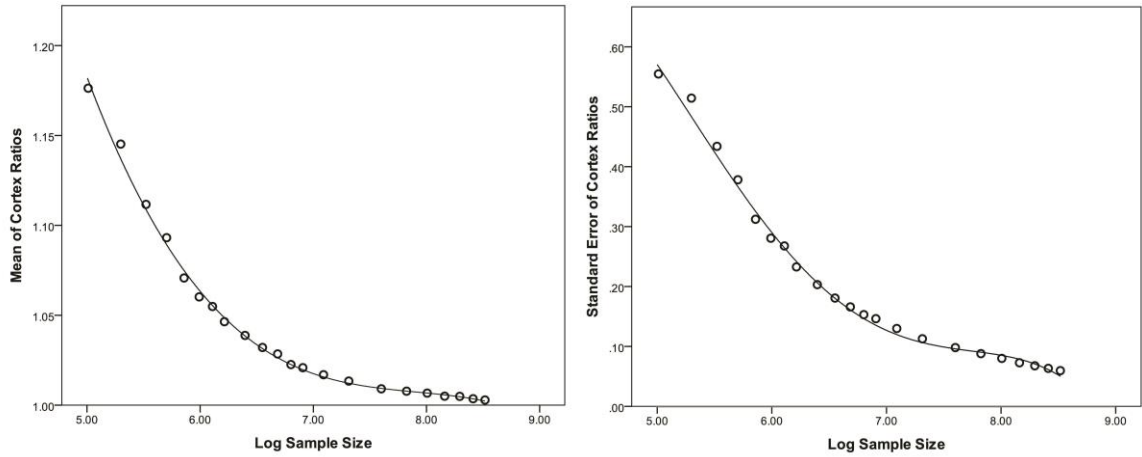
Figure 7 – Scatter plots of the mean and standard error of Cortex Ratios from the sampling distributions generated from the experimental dataset with varying sample sizes (log-transformed). Quadratic curve models are fitted to the two plots.

The following two equations summarize the polynomial models fitted to the two plots:

$$\mu=-.00054248x^4+.0082668x^3+.0048148x^2-.55961x+3.16565 \ (n=23; \ r=.998)$$

where μ is the mean Cortex Ratio of the sampling distribution and x is the log transformed sample size.

$$se=-.0091008 \ x^4+ .23755x^3- 2.24241x^2+ 8.87460x- 11.74312 \ (n=23; \ r=.998)$$

where se is the standard error of the Cortex Ratio sampling distribution and x is the log transformed sample size. By using the above two equations and substituting x with the size of archaeological assemblages, it is possible to compute the confidence interval of Cortex Ratios for any given assemblage. Clearly, this method requires the assumption that Cortex Ratios are normally distributed. As pointed out before, Cortex Ratio sampling distributions are somewhat skewed when the sample size is small, although the distribution does quickly become symmetrical as sample size increases. Table 6 compares the significance levels of the study assemblages determined by the 95% confidence intervals based on the above two equations and the probabilities obtained from the Monte Carlo simulations presented earlier. The results of the two approaches show general agreements except for a few layers. At this point, it is unclear why the disagreements exist, although factors such as sample size likely play a key role in affecting the resulting significance levels. Thus, while the confidence interval approach may be useful to provide preliminary statistical information regarding archaeological Cortex Ratios, more thorough statistical treatment, such as the Monte Carlo sampling technique employed here, is still needed for more concrete assignments of statistical probability.

148

Layer	Cortex Ratio	Assemblage Size	Monte Carlo p (from assemblage with ratio of 1)*	p (95% Confidence Interval)*
Roc de Marsal				
2	.71	685	**.019**	>.05 (.67-1.40)
3	.81	174	.43	>.05 (.12-2.19)
4	.68	2485	**<.0001**	**<.05 (.83-1.19)**
5	.83	1780	**.042**	>.05 (.81-1.21)
6	.94	501	.70	>.05 (.57-1.53)
7	.92	2758	.28	>.05 (.83-1.18)
8	.97	2410	.71	>.05 (.83-1.19)
9	1.02	4651	.75	>.05 (.89-1.12)
Pech de l'Azé IV				
3A	1.28	1547	**.042**	**<.05 (.80-1.22)**
3B	1.19	2762	**.041**	**<.05 (.83-1.18)**
4A	.87	216	.57	>.05 (.22-2.04)
4B+C	1.05	647	.80	>.05 (.65-1.42)
5A	1.25	1366	.076	**<.05 (.79-1.23)**
5B	1.00	622	.98	>.05 (.64-1.43)
6A	1.01	2309	.92	>.05 (.83-1.19)
6B	1.10	2304	.31	>.05 (.83-1.19)
7	.83	2784	**.015**	**<.05 (.83-1.18)**
8	1.01	2076	.95	>.05 (.82-1.20)
Combe-Capelle Bas				
I-upper	.60	2167	**<.0001**	**<.05 (.82-1.20)**
I-lower	.86	2222	.071	>.05 (.82-1.19)
I-3	.76	109	.43	>.05 (.0-2.52)
II-3	.75	207	.27	>.05 (.20-2.07)
II-4	.73	1484	**.0012**	**<.05 (.80-1.22)**
III-1	.46	525	**<.0001**	**<.05 (.59-1.51)**

*Two-tailed probability

Table 6 – Comparison of statistical significance levels derived from the experimental data through Monte Carlo sampling and using 95% confidence interval based on the regression models. Bold values indicate significance level exceeds an alpha of .05. Bracketed values denote the range of the confidence interval.

The second challenge of the cortex methodology concerns the ability to compare two Cortex Ratios in a statistically confident manner. In this study, a permutation test is employed to construct the null distributions of the differences between resampled Cortex Ratios, with which the actual difference between archaeological ratios can be compared. Results from both tests have clear implications for the analytical potential of small assemblages with sizes in the 100-200 range, where their Cortex Ratios are difficult to compare and interpret due to the high degree of variance. This difficulty arises because the sampling distribution of Cortex Ratios from smaller assemblages is much wider, which therefore increases the chance for the difference between two archaeological ratio values to occur in frequencies over the alpha threshold. This observation echoes concerns over the effects of assemblage size on richness and diversity (Grayson and Cole, 1998; Hiscock, 2001; Meltzer et al., 1992), and likely extends to other properties, such as flake to core ratio, tool to flake ratio, and flake percentage that are commonly used for describing assemblages, and therefore warrants further examination. A better grasp of the potential impact of sample size on Cortex Ratios and other assemblage measures would allow us to investigate different archaeological samples more meaningfully.

Roc de Marsal and Pech de l'Azé IV show similar cortex patterns in the lower layers but opposite patterns in upper layers. Statistical evaluation demonstrate that the trend at Roc de Marsal was more gradual whereas that at Pech de l'Azé IV was more punctuated and fluctuating shifts between layers. What is clear, however, is that most assemblages which correspond to colder and more open environments exhibit rather extreme cortex proportions (Layers 2 and 4 at Roc de Marsal, Layers 3A, 3B, and 5A at

Pech de l'Azé IV, all layers at Combe-Capelle Bas) compared to layers that contain

faunal records characteristic of temperate settings. The pattern observed here suggests

that there was little artifact movement during warmer periods that altered the assemblage

cortex and volume composition at these localities. By contrast, during colder periods,

greater degrees of artifact transport occurred and caused the apparent imbalance in

assemblage cortex composition. Specifically, patterns of artifact movement during this

period led to a gradual loss of cortex at Roc de Marsal while Pech de l'Azé IV saw

greater fluctuation and excess in cortex proportions. At Combe-Capelle Bas, the layers

exhibited an overall trend of underrepresented cortex. The considerable variability in

Cortex Ratios between the stratigraphic units might also suggest the underlying

behavioral processes that gave rise to these patterns were likely temporally varied and

complex.

The results here fit with the model of Middle Paleolithic mobility strategies

proposed by Delagnes and Rendu (2011). At Roc de Marsal and Pech de l'Azé IV, the

Levallois rich layers at the lower half of the stratigraphic sequence exhibit less extensive

artifact movement and therefore likely reflect lower degrees of mobility. On the other

hand, non-Levallois industries, such as the Quina and MTA technology, dominate layers

with Cortex Ratios showing greater deviation from 1 at all three sites. This trend likely

indicates a shift in technological strategies geared towards heightened mobility pattern

perhaps in response to increasing environmental unpredictability (Delagnes and Rendu

2011; Hiscock et al. 2009; Niven et al. 2012). However, departures in the Cortex Ratios

between the three sites suggest variation in the nature and extent of artifact transport between places even within a single technological system or climatic setting.

At this point, it is worth considering the sort of artifact transport patterns that could have led to the formation of the observed differences in Cortex Ratios. A Cortex Ratio of 1 may suggest that no artifact import or export had occurred at a locality and therefore the unity between cortex and stone volume remained intact. On the other hand, a transport pattern based purely on random selection would also not affect the cortex to volume ratio of an assemblage. This scenario is nevertheless unlikely as it would require all artifacts regardless of size and type to have the same probability of being transported, which contradicts much of the discussion on artifact utility where size and cutting edge has been identified as a central criterion in terms of functionality, transportability, and utility (Kuhn 1994; Lin et al. 2013; Shott and Sillitoe 2005).

The other possibility consists of repeated back and forth movements of artifacts between places through a more tortuous (non-linear) mobility pattern that causes assemblage cortex compositions to remain relatively balanced overtime (Douglass 2010). This kind of movement would involve less long linear movements and more frequent "turns" (Brantingham 2006). The consistent pattern of cortex to volume at unity represented in the lower layers at Roc de Marsal and Pech de l'Azé IV opens the possibility that this sort of artifact movement pattern took place at the two localities during this period. If we consider Roc de Marsal and Pech de l'Azé IV as places of not only artifact production but also of discard, reuse, and retooling, then the cortex pattern

can be explained by the recurring production, selection, import, and export of artifacts at these sites. In other words, there was potentially great overlap in the utilization of these places by Neanderthals during warmer and forested conditions.

For the following colder periods, potentially corresponding to MIS 3 and 4, the greater excess and deficit in cortex at the three sites indicate a tendency for artifacts to be transported away from the sites and not come back as frequently, or vice versa. This pattern suggests a changing Neanderthal mobility pattern possibly between unevenly distributed and spatially localized resource patches (Delagnes and Rendu, 2011; Kuhn, 1995). The distinct cortex patterns during this period at Roc de Marsal and Pech de l'Azé IV also suggest that forms of artifact transport differ spatially between sites (see Soressi 2002). The lithic technologies associated with these layers, including the Quina and MTA, have been characterized as having greater emphasis on the renewal and maintenance of artifact utility, especially through resharpening (Delagnes and Rendu, 2011; Hiscock et al., 2009; Turq 1989; Niven et al., 2012). These observations point to the tendency for Neanderthals to carry out long distance movements during this period. Such movement patterns may relate to the targeted exploitation of migratory reindeer during colder periods where the ecological conditions was more heterogeneous (Delagnes and Rendu 2011).

However, the nature of the underlying processes that gave rise to this difference in assemblage cortex proportions remains unclear, especially given that specific cortex ratios can be created by multiple assemblage configurations. The deficit of cortex at Roc

de Marsal can be explained by either an export of cortex or import of volume and vice versa for the excess of cortex at Pech de l'Azé IV. Furthermore, the relationship between Cortex Ratio variability and the various facets of occupation and mobility, such as the regularity and duration of (re)occupation or the frequency, velocity, and linearity of movement, need to be further assessed. Further examination of the relationship between Cortex Ratios and other technological variables are required to provide a more contextual understanding of the underlying behavioral process. It is also important to note that the assemblages considered in this study only include those that are attributed to locally derived flint types. Although these elements compose the majority of the study assemblages, the analysis here only partially captures the artifact movement patterns represented in these three sites. Additional studies of assemblage materials made on other raw material types are needed to shed light on other aspects of Middle Paleolithic artifact movement patterns.

4.8 Conclusion

Even in the context where workable materials are abundant, stone artifacts were still transported as populations moved and carried out activities across the landscape (Douglass and Holdaway 2011; Douglass 2010; Douglass et al. 2008). In other words, the use of stone artifacts was likely more widespread than the localized places where raw materials occur. As Douglass and Holdaway (2011) have stated, from a landscape perspective there was likely always a tendency for stone artifacts to be moved with varying distances from some localities for use and discard elsewhere. A similar observation was recently made by Turq et al. (2013) concerning the movement of raw

material in the Middle Paleolithic of Western Europe. This interaction between movement and the formation of lithic assemblages across different places over the landscape means that mobility can be conceptualized by the over or under abundance of knapping products at particular localities (Dibble et al. 2005; Douglass and Holdaway 2011; Holdaway et al. 2008).

From this perspective, cortex quantification offers a viable alternative for gauging the nature and extent of past mobility and land use patterns. The ability to establish statistical confidence between Cortex Ratios thus further extends the approach's analytical utility to archaeological materials. In addition, it is important to note that Cortex Ratios merely summarize the cortex composition of lithic assemblages and do not offer immediate behavioral explanations. In particular, the impact of various contextual factors, including cobble size and reduction intensity, need to be taken into consideration. Future work is needed to further examine and test the nuanced interaction between Cortex Ratios and various forms of artifact movement and behavioral processes.

CHAPTER 5: Summary and Conclusions

"Science seeks to build a cumulative learning trajectory through the study of its empirical subject matter" (Binford, 2001:672).

The line of research presented in this dissertation explores the application of an archaeological science perspective to the study of prehistoric stone artifacts. The rationale for this body of research stems from the inferential challenges faced in current stone artifact archaeology, especially the shift to lower levels of inferential confidence as researchers seek to address higher order topics of hominin behavior and adaptation. This issue of inferential adequateness and validity is attributed to the common conflation of behavioral assumptions within basic archaeological units and the dependence on simplistic analogic treatment of uniformitarian theory and experimental design. Alternatively, by approaching lithic studies as an archaeological science, it is argued here that stone artifact archaeology can develop as a productive scientific enterprise in its own right by applying greater rigor to the design of experimental research and the overall process of archaeological inference building.

Chapter 2 provided an analysis on the role of analogy and experimentation in the building of lithic inference. The study described therein first critically examined the nature of analogic reasoning in archaeology, and related the inferential issues in much of archaeological systematics to the operation of formal analogy. In stone artifact archaeology, conventional replicative experimentation lacks elements of experimental

rigor, causing its reasoning process to fall in line with formal analogy. Consequently, the generated inferences are prone to issues of analogical inadequacy.

To establish sound relational inferences of high validity between lithic artifacts and past causal processes, it is argued that rigorous treatment of experimental design is necessary in terms of variable control and hypothesis formulation. The focus of lithic experimentation should also shift from replication of artifacts to a comparative approach based on the evaluation of archaeological assemblages with reference to experimentally verified relationships attributable to controlled variables. The chapter advocated a two-step procedure to lithic experimentation – 'pilot' experiments that obtain a general understanding of the phenomenon in question through an emphasis on high ecological validity; and, 'second generation' experiments that explicitly test the stipulated relationships under controlled settings to ensure high internal validity of the experimental outcome. The repeated back-and-forth between pilot and second generation experiments, along with continual comparative feedback from the archaeological record, allows concrete linkages of referential knowledge to be established in a constructive manner for drawing broader archaeological inference.

The following two chapters investigated the use of the proposed comparative experimental approach on the study of stone artifacts. **Chapter 3** employed a highly controlled experimental setup to examine uniformitarian linkages between platform configurations and the formation of lithic attributes that are relatable to higher level technological behaviors. Specifically, this paper explored how particular characteristics

157

of individual, unretouched flakes can be altered in ways that increase their relative utility and economy, as reflected in the ratio of flake edge length to mass. Experimental results identified exterior platform angle and platform depth as being primary independent variables affecting this ratio. These relationships were tested against a number of Middle Paleolithic archaeological assemblages with distinct manufacturing characteristics. The results indicated diachronic patterns in flake utility and economization, possibly related to differences in tool-kit selection and maintenance associated with changes in ranging and provisioning behavior of late Pleistocene Neanderthals in southwestern France.

Chapter 4 applied the cortex quantification approach to the lithic assemblages of three Middle Paleolithic sites in southwestern France. The geometric method quantified the cortex proportion among assemblages as the Cortex Ratio, which indicates the excess or deficit of assemblage cortex relative to existing artifact volume. This paper determined statistical confidence for archaeological Cortex Ratios through Monte Carlo and resampling techniques. Experimentally produced data generated from local raw material samples with known geometric properties and measured cortex values were employed as a reference for attaching a probability to an archaeological assemblage's Cortex Ratio. The archaeological results indicated changes in assemblage cortex proportion among the study assemblages over time, namely a tendency for assemblages associated with colder, dryer, and more open environments to exhibit Cortex Ratios that deviate from a value of one. This observation was discussed with respect to the possible artifact movement patterns that could have contributed to the formation of the observed assemblage cortex composition.

158

The results of Chapter 3 and 4 revealed interesting diachronic patterns in the lithic assemblages from the three Middle Paleolithic sites studied here. Specifically, a clear association is demonstrated between lithic assemblages characterized by (1) small, thin flakes with high cutting-edge-to-mass ratio, (2) relatively fewer scrapers, and (3) values of assemblage volume and cortex that approach unity. These assemblages are further linked to temperate forest environmental conditions as indicated by the dominance of red deer and roe deer remains. The close or equal-to-one Cortex Ratios of these assemblages indicate that the selection and movement of artifacts occurred in ways that did not cause imbalance in the original assemblage cortex composition. These findings suggest that either the selection of artifacts was based on criteria unrelated to geometry and hence utility (as defined by the relative amount of usable edge), or the movement of artifacts between places was more 'tortuous' (i.e., non-linear path) and caused assemblage cortex composition to remain relatively balanced overtime (Douglass 2010). This kind of movement would involve fewer long linear movements and more frequent "turns" that result in more thorough use of a localized landscape (Brantingham 2006).

Based on models of technological organization and risk management (Elston 1990; Nelson 1991), this ranging strategy would involve less frequent movements where extensive tool-kit provisioning and curation is required. This scenario is supported by the general abundance of small and thin flakes in these assemblages that shows an emphasis on the production of immediately usable cutting edge. The proportionally lower frequencies of retouched scrapers among these assemblages may relate to either an overall lower rate of flake utility rejuvenation (Lin et al. 2013), or that the activities

involving artifact use-life maintenance occurred away from these cave localities and over the wider river valley.

In contrast, some other assemblages, particularly those from Roc de Marsal, are characterized by (1) larger and thicker flakes, (2) higher proportions of scrapers, (3) Cortex Ratios that statistically differ from a value of one, and (4) a tendency to be associated with colder and dryer environments, as indicated by the dominance of reindeer remains in these assemblages. The Cortex Ratios suggest that the nature and degree of artifact movement likely occurred in ways that caused an imbalance in the assemblage cortex proportions – i.e., artifacts moved away from places but did not come back as frequently, or vice versa. This pattern may relate to ranging behaviors consisting of long linear distant movements away from places (Brantingham 2006; Douglass 2010). The dominance of large, thick flakes and the higher proportions of retouched scrapers among these assemblages further suggest that focus of tool-kit provisioning were placed on the selective transport of large flakes with greater potential for utility rejuvenation (Lin et al. 2013).

This shift in Neanderthal ranging and tool-provisioning behavior likely reflects broader changes in land use. Specifically, the transition from a warmer and more temperate climate to a colder and dryer one, likely corresponding to the move from Marine Isotope Stage 5 to 4, would have had an effect on the overall distribution and abundance of resources, specifically large mammals. Isotopic studies have shown that the Neanderthal diet mainly consisted of large herbivores, including large bovids, horses,

reindeer, wooly rhinoceros, mammoth, and large deer (Bocherens 2009; Bocherens et al. 2005; Richards and Trinkaus 2009; Richards et al. 2000, 2008; Salazar-García et al. 2013). In long term cold climatic episodes, permafrost conditions and the decline in vegetation cover would have increased surface sediment runoff due to greater overland fluvial discharge (Bogaart et al. 2003; Vandenberghe 1995, 2002, 2003). This process would have led drainage organization to change from a single and stable river system (e.g., meandering) to a more ephemeral and multi-channeled system (e.g., braided or anabranching) (Bertran et al. 2013; Vandenberghe 2001, 2008). Under this scenario, the occurrence of large herbivores during cold pleniglacial episodes could have fluctuated with greater spatial and temporal heterogeneity in the Dordogne region due to seasonal changes in snow/frost cover. In conjunction with the lower vegetation diversity compared to warmer interglacial cycles (Rivals et al. 2009), the observed archaeological shifts may reflect Neanderthal foraging responses to major alterations in food resource composition and distribution across the landscape.

However, as emphasized in the beginning of this dissertation, much more research into a variety of topics will be required to confidently and holistically arrive at a higher level inference of late Pleistocene Neanderthal land use across the Dordogne landscape. In particular, the current study sample is small and restricted to cave localities. It is therefore unclear whether the observed archaeological pattern is representative of wider behavioral processes, or only captures a subset of past Neanderthal dynamics limited to these cave settings. Furthermore, a greater understanding of paleoenvironmental

conditions of the region would allow a more in-depth assessment of foraging and land use models in relation to changes in prey distribution and movement.

Related to this point, more contextual information about the changes in raw material availability (i.e., distribution in primary and secondary context, visibility, abundance, morphology, etc.) between the two climatic periods is also critical in discerning the factors affecting the shift in flake production strategies and conditions for artifact selection and transport. In addition, application of the same analytical procedure to Upper Paleolithic assemblages in the same region would provide a useful comparison for evaluating the differences in behavioral patterns between Neanderthals and modern humans under similar ecological settings. This sort of dataset would prove invaluable to the understanding of the differences between the behavioral capacities of Neanderthals and Upper Paleolithic modern human populations.

Finally, while it is advocated here that archaeological units of measurement should be established upon uniformitarian processes, behavioral interpretations for explaining the observed patterns cannot be sought in a hypo-deductive manner. This is because such reasoning exercise necessitates a priori assumptions about hominin behavior. In doing so, we would fall back into the cycle of inferential fallacy of formal analogy. Instead, behavioral interpretations may be more effectively constructed as inference to the best explanation (Fogelin 2007). Contrary to the deduction-induction archaeological hermeneutics (Hodder 1999; Wandsnider 2004), this reasoning process is abductive in nature, being based on the assumption that the explanation capable of

162

accounting for the most evidence is also the most likely to be true (also see Wylie 1993).

By increasing the number of lines of independent evidence, each grounded in sound

inferential framework, it is thus possible to construct complex explanations by testing the

relative capacity of competing explanatory models, possibly through approaches such as

simulation and agent-based modeling.

Returning to the broader discussion of science and lithic studies, a common

definition of archaeological science is the application of scientific techniques to the study

of archaeological materials (Tite 1991). However, exactly what makes a technique

'scientific' has yet to be critically evaluated and defined. More recently, the definition

seems to have been made explicit by specifying the use of computational modeling and

natural science methods for anthropological and archaeological inquiry. This view is

increasingly apparent in current stone artifact archaeology. Advanced computational

techniques such as 3D digitization and geometric-morphometrics are now commonly

featured in the analysis of prehistoric stone artifacts (Bretzke and Conard 2012; Clarkson

2010; Iovita and McPherron 2011; Iovita 2009, 2011; Lin et al. 2010; Lycett and von

Cramon-Taubadel 2013; Lycett et al. 2010; Rezek et al. 2011). Renewed attention to

quantification, objectification, modelling, and hypothesis testing have equally been

revived recently (Lycett and Chauhan 2010) in conjunction with the growing application

of outside frameworks, including phylogenetic and transmission models to lithic analysis

(Lycett 2009a,b; O'Brien et al. 2001).

However, while studies embrace this movement towards the adoption of more sophisticated and rigorous techniques, most of which are developed in other fields, the underlying theoretical framework that governs the inferential logic of stone artifact archaeology remains largely unchanged. Artifacts continue to be treated as emic units, from which past intentionality, decisions, goals, and cultural affinity are reconstructed via replicative experiments on the basis of modern observers' ordinary experience, common sense, and intuition. At one level, this approach can be attributed to a general absence of scientific training in archaeology programs, which results in the implicit notion that mere quantification and experimentation equal science. In the absence of developed theory and understanding of process, the simple pursuit of quantification and objectification has been shown to be largely unproductive in previous attempts (Dunnell 1982). In addition, while experimentation has proved to be an increasingly popular approach, the lack of attention to its nature and role has led to many of the generated inferences being difficult to compare and evaluate.

In some ways, this issue is reflected by the need for Dibble and Whittaker (1981) to ironically emphasize the 'control' aspect of lithic experimentation, given that the ability to manipulate and control variables is commonly viewed as the chief virtue of experimentation as a scientific method (Gauch 2003; Kosso 2011). At a higher level, the recurring theme of concept and technique borrowing from outside disciplines in archaeology (Dunnell 1982; Hodder 2012) points to a general deficiency of archaeological theory to deal with the ontological nature of the archaeological record and its inferential linkage to behavioral and cultural interpretations (Binford 2001; Clarke

164

1968). The apparent contradictory, yet interchangeable ways stone artifacts are currently treated (e.g., units of cultural affinity, functional design, chronological pattern, continuous reduction, or natural selection), signal that the nature and role of theory in lithic studies is in disarray (see Schiffer, 1988).

If we take the notion of scientific practice as outlined in the opening quote to this chapter by Binford (2001), science entails having a clear and explicit understanding of its subject matter and the ability to produce meaningful inferences based on solid inferential grounding. More importantly, the scientific process of inference building is purposefully slowed down in order to establish sound control over variables and analytical transparency (Kosso 2011). The scientific method is powerful as it can take something that is obvious to our ordinary experience and show its actual complicated nature. Indeed, from radiometric and geochronological dating, to isotopic analysis and paleoenvironmental reconstructions, practices of archaeological science have provided incredible insights and discoveries into the past that are simply beyond the scope and imagination of ordinary researchers.

It has been argued that the role of science in archaeology does not exist as an unified enterprise, and the credibility and utility of archaeological investigation is related more to the wide range of theories and techniques that make interdisciplinary interactions possible and effective (Wylie 2000). Indeed, archaeology is situated in the unique position of being able to utilize and integrate knowledge from various disciplines to study the past. However, as demonstrated throughout this dissertation, it is equally important to

165

be critically aware of the goals and the associated inferential reasoning processes that underlie such an interdisciplinary effort. As David Clarke (1968:13) famously expressed, "Archaeology is archaeology is archaeology". Few would disagree that archaeological inquiry is grounded in material reality, and that this material reality – the archaeological record – constitutes the primary subject of study for archaeologists. Given the potential complex formational history of the record, as well as issues such as equifinality and time scale difference, it cannot be assumed that this record can be intuitively translated into units of behavior based on the common sense or daily ordinary experience of the observers. It is, thus, important for archaeologists to construct past interpretations on the basis of a thorough understanding of its study subject.

The point here is that the aspiration of scientific practice emphasizes the ontological and formational component of archaeological inquiry – i.e., what is the material reality of the archaeological record, and how was it formed in relation to past behavioral processes? It is argued here that the virtue and analytical power of archaeological science resides in the use of principles from the natural sciences as the foundation for building archaeological interpretations. Because of their uniformitarian nature and, more importantly, their independence from the anthropological phenomenon that we wish to investigate, these principles serve as powerful tools for establishing referential foundations and linkages for constructing archaeological inferences.

As this dissertation has demonstrated, the application of a similar scientific design and a methodological treatment of uniformitarian theory to lithic studies do not

necessarily restrict lithic research to the discovery of low level physical phenomena. Indeed, the disconnect between the mundane material reality of archaeological objects and the behavioral dynamics of past human phenomena that archaeologists seek to understand is often translated into a dilemma between inferential security and interpretive details (Wylie 1985). Earlier attempts to search for human universals have instead resulted in behaviors that are either trivial or too obvious to be mentioned (Flannery 1973).

This dissertation outlines an alternative strategy for resolving this issue by suggesting a move away from the search for singular causal connections for linking archaeological material to past human processes that we wish to evaluate. More specifically, research cannot begin with the assertion that there is an inherent and meaningful uniformitarian connection between the subject of study and the research question, and that this inferential connection can somehow be intuitively discovered. Instead, because all archaeological phenomena are grounded in material reality and therefore obey universal laws, the approach advocated here involves the use of uniformitarian processes to detect physical patterns in stone artifacts and draw behavioral inferences about the possible underlying formational processes.

BIBLIOGRAPHY

Abdi, H., Edelman, B., Valentin, D. and Dowling, W. J. (2009). *Experimental Design and Analysis for Psychology*, Oxford University Press, Oxford.

Ahler, S. A. and Geib, P. R. (2000). Why Flute? Folsom Point Design and Adaptation. *Journal of Archaeological Science* **27**: 799–820.

Ambrose, S. H. (2001). Paleolithic Technology and Human Evolution. *Science* **291**: 1748–1753.

Amick, D. S., Mauldin, R. P. and Binford, L. R. (1989). The Potential of Experiments in Lithic Technology. In D. S. Amick and R. P. Mauldin (eds.), *Experiments in Lithic Technology*, Archaeopress, Oxford, pp.1–14.

Amick, D. S., Mauldin, R. P. and Tomka, S. A. (1988). An Evaluation Of Debitage Produced By Experimental Bifacial Core Reduction of a Georgetown Chert Nodule. *Lithic Technology* **17**: 26–36.

Andrefsky, W. (1994). Raw-Material Availability and the Organization of Technology. *American Antiquity* **59**: 21–34.

Andrefsky, W. (2005). *Lithics: Macroscopic Approaches to Analysis*, Cambridge University Press, Cambridge.

Andrefsky, W. (2006). Experimental and Archaeological Verification of An Index of Retouch for Hafted Bifaces. *American Antiquity* **71**: 743–757.

Andrefsky, W. (2009). The Analysis of Stone Tool Procurement, Production, and Maintenance. *Journal of Archaeological Research* **17**: 65–103.

Archer, W. and Braun, D. R. (2010). Variability in Bifacial Technology at Elandsfontein, Western cape, South Africa: A Geometric Morphometric Approach. *Journal of Archaeological Science* **37**: 201–209.

Ascher, R. (1961a). Analogy in Archaeological Interpretation. *Southwestern Journal of Anthropology* **17**: 317–325.

Ascher, R. (1961b). Experimental Archeology. *American Anthropologist* **63**: 793–816.

Bailey, G. (1983). Concepts of Time in Quaternary Prehistory. *Annual Review of Anthropology* **12**: 165–192.

Bamforth, D. B. (1986). Technological Efficiency and Tool Curation. *American Antiquity* **51**: 38–50.

Bamforth, D. B. (1990). Settlement, Raw Material, and Lithic Procurement in the Central Mojave Desert. *Journal of Anthropological Archaeology* **9**: 70–104.

Bamforth, D. B. (1991). Technological Organization and Hunter-Gatherer Land Use: A California Example. *American Antiquity* **56**: 216–234.

Bamforth, D. B. and Bleed, P. (1997). Technology, Flaked Stone Technology, and Risk. *Archaeological Papers of the American Anthropological Association* **7**: 109–139.

Barham, L. S. (1987). The Bipolar Technique in Southern Africa: A Replication Experiment. *The South African Archaeological Bulletin* **42**: 45–50.

Bartovics, A. F. (1974). The Experiment in Archaeology: A Comparison of Two Case Studies. *Journal of Field Archaeology* **1**: 197–205.

Bar-Yosef, O. and Dibble, H. L. (1995). Preface. In H. L. Dibble and O. Bar-Yosef (eds.), *The Definition and Interpretation of Levallois Technology*, Prehistory Press, Madison, pp.ix–xiii.

Bar-Yosef, O. and Kuhn, S. L. (1999). The Big Deal about Blades: Laminar Technologies and Human Evolution. *American Anthropologist* **101**: 322–338.

Bar-Yosef, O. and Van Peer, P. (2009). The Chaîne Opératoire Approach in Middle Paleolithic Archaeology. *Current Anthropology* **50**: 103–131.

Bertran, P., Frouin, M., Mercier, N., Naessens, F., Prodeo, F., Queffelec, A., Sirieix, C. and Sitzia, L. (2013). Architecture of the Lower Terraces and Evolution of the Dordogne River at Bergerac (South-West France) During the Last Glacial-Interglacial Cycle. *Journal of Quaternary Science* **28**: 605–616.

Bettinger, R. L. (1987). Archaeological Approaches to Hunter-Gatherers. *Annual Review of Anthropology* **16**: 121–142.

Binford, L. R. (1962). Archaeology as Anthropology. *American Antiquity* **28**: 217–225.

Binford, L. R. (1965). Archaeological Systematics and the Study of Culture Process. *American Antiquity* **31**: 203–210.

Binford, L. R. (1967). Smudge Pits and Hide Smoking: The Use of Analogy in Archaeological Reasoning. *American Antiquity* **32**: 1–12.

Binford, L. R. (1973). Interassemblage Variability: the Mousterian and the "Functional" Argument. In C. Renfrew (ed.), *The Explanation of Culture Change*, Duckworth, London, pp.227–254.

Binford, L. R. (1977a). Forty-Seven Trips: A Case Study in the Character of Archaeological Formation Processes. In R. S. V Wright (ed.), *Stone Tools as Cultural Markers*, Australian Institute of Aboriginal Studies, Canberra, pp.24–36.

Binford, L. R. (1977b). General introduction. In L. R. Binford (ed.), *For Theory Building in Archaeology*, Academic Press, New York, pp.1–10.

Binford, L. R. (1978a). Dimensional Analysis of Behavior and Site Structure: Learning from an Eskimo Hunting Stand. *American Antiquity* **43**: 330–361.

Binford, L. R. (1978b). *Nunamiut Ethnoarchaeology*, Academic Press, New York.

Binford, L. R. (1979). Organization and Formation Processes: Looking at Curated Technologies. *Journal of Anthropological Research* **35**: 255–273.

Binford, L. R. (1980). Willow Smoke and Dogs' Tails: Hunter-Gatherer Settlement Systems and Archaeological Site Formation. *American Antiquity* **45**: 4–20.

Binford, L. R. (1981). *Bones: Ancient Men and Modem Myths*, Academic Press, New York.

Binford, L. R. (1987). Data, Relativism and Archaeological Science. *Man, New Series* **22**: 391–404.

Binford, L. R. (2001). Where Do Research Problems Come From? *American Antiquity* **66**: 669–678.

Binford, L. R. (2002). *In Pursuit of the Past: Decoding the Archaeological Record*, University of California Press, Berkeley.

Binford, L. R. and O'Connell, J. F. (1984). An Alyawara Day : The Stone Quarry. *Journal of Anthropological Research* **40**: 406–432.

Blades, B. S. (2003). End Scraper Reduction and Hunter-Gatherer Mobility. *American Antiquity* **68**: 141–156.

Bleed, P. (1986). The Optimal Design of Hunting Weapons: Maintainability or Reliability. *American Antiquity* **51**: 737–747.

Bleed, P. (1996). Risk and Cost in Japanese Microcore Technology. *Lithic technology* **21**: 95–107.

Bleed, P. (2001). Trees or Chains, Links or Branches: Conceptual Alternatives for Consideration of Stone Tool Production and Other Sequential Activities. *Journal of Archaeological Method and Theory* **8**: 101–127.

Bocherens, H. (2009). Neanderthal Dietary Habits: Review of the Isotopic Evidence. In J.-J. Hublin and M. P. Richards (eds.), *The Evolution of Hominin Diets: Integrating Approaches to the Study of Palaeolithic Subsistence*, Springer, New York, pp.241–250.

Bocherens, H., Drucker, D. G., Billiou, D., Patou-Mathis, M. and Vandermeersch, B. (2005). Isotopic Evidence for Diet and Subsistence Pattern of the Saint-Césaire I Neanderthal: Review and Use of A Multi-Source Mixing Model. *Journal of Human Evolution* **49**: 71–87.

Boëda, É. (1986). Approche Technologique du Concept Levallois et Evaluation de son Champ d'Application: Étude de Trois Gisements Saaliens et Weichseliens de la France Septentionale, Unpublished Ph.D. thesis, Université de Paris IX.

Boëda, É. (1988). Le Concept Laminaire: Rupture et Filiation avec le Concept Levallois. In J. K. Kozlowski (ed.), *L'homme de Neandertal: Actes Du Colloque International de Liege, Vol. 8, La Mutation*, Universite de Liege, Liege, pp.41–59.

Boëda, É. (1993). Le Débitage Discoide et le Débitage Levalllois Recurrent Centripète. *Préhistorique Française* **90**: 392–404.

Boëda, É. (1994). *Le Concept Levallois: Variabilité des Methodés*, CNRS, Paris.

Boëda, É. (1995). Levallois: A Volumetric Construction, Methods, A Technique. In O. Bar-Yosef and H. L. Dibble (eds.), *The Definition and Interpretation of Levallois Technology*, Academic Press, New York, pp.41–68.

Boëda, É., Bonilauri, S., Connan, J., Jarvie, D., Mercier, N., Tobey, M., Valladas, H., al Sakhel, H. and Muhesen, S. (2008). Middle Palaeolithic Bitumen Use at Umm el Tlel around 70,000 BP. *Antiquity* **82**: 853–861.

Boëda, É., Connan, J., Dessort, D., Muhesen, S., Mercier, N., Valladas, H. and Tisnerat, N. (1996). Bitumen as A Hafting Material on Middle Palaeolithic Artefacts. *Nature* **380**: 336–338.

Boëda, É., Geneste, J.-M. and Meignen, L. (1990). Identification de Chaîne Opératoire Lithiques du Paléolithique Ancien et Moyen. *Paléo* **2**: 43–80.

Boëda, É., Hou, Y.-M., Forestier, H., Sarel, J. and Wang, H. M. (2013). Levallois and Non-Levallois Blade Production at Shuidonggou in Ningxia, North China. *Quaternary International* **295**: 191–203.

Bogaart, P. W., Tucker, G. E. and de Vries, J. J. (2003). Channel Network Morphology and Sediment Dynamics Under Alternating Periglacial and Temperate Regimes: A Numerical Simulation Study. *Geomorphology* **54**: 257–277.

Bonnichsen, R. (1977). *Models for Deriving Cultural Information from Stone Tools*, National Museums of Canada, Ottawa.

Bordes, F. (1961). *Typologie du Paleolithique Ancien et Moyen*, Delmas, Bordeaux.

Bordes, F. (1976). Le Gisement de Pech de l'Azé IV. Note Préliminaire. *Bulletin de la Société Préhistorique Française* **72**: 293–308.

Bourguignon, L. (1997). *Le Moustérien de Type Quina: Nouvelle Définition d'une Entité Technique*, unpublished Ph.D. thesis, Université de Paris X.

Bousman, C. B. (1993). Hunter-Gatherer Adaptations, Economic Risk and Tool Design. *Lithic Technology* **18**: 59–86.

Bousman, C. B. (2005). Coping with Risk: Later Stone Age Technological Strategies at Blydefontein Rock Shelter, South Africa. *Journal of Anthropological Archaeology* **24**: 193–226.

Bouzouggar, A. (1997). *Matières Premières, Processus de Fabrication et de Gestion des Supports d'Outils dans la Séquence Atérienne de la Grotte des Contrebandiers à Temara*, unpublished Ph.D. thesis, Université de Bordeaux.

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

Bradbury, A. P. and Carr, P. J. (1999). Examining Stage and Continuum Models of Flake Debris Analysis: An Experimental Approach. *Journal of Archaeological Science* **26**: 105–116.

Brantingham, P. J. (2006). Measuring Forager Mobility. *Current Anthropology* **47**: 435–459.

Brantingham, P. J. (2010). The Mathematics of Chaînes Opératoires. In S. Lycett and P. R. Chauhan (eds.), *New Perspectives on Old Stones: Analytical Approaches in Paleolithic Technologies*, Springer, New York, pp.183–206.

Brantingham, P. J. and Kuhn, S. L. (2001). Constraints on Levallois Core Technology: A Mathematical Model. *Journal of Archaeological Science* **28**: 747–761.

Brantingham, P. J., Olsen, J. W., Rech, J. A. and Krivoshapkin, A. I. (2000). Raw Material Quality and Prepared Core Technologies in Northeast Asia. *Journal of Archaeological Science* **27**: 255–271.

Braun, D. R. (2006). *Ecology of Oldowan Technology: Koobi Fora and Kanjera South*, unpublished Ph.D. dissertation, Rutgers University.

Braun, D. R., Harris, J. W. K. and Maina, D. N. (2009a). Oldowan Raw Material Procurement and Use: Evidence from the Koobi Fora Formation. *Archaeometry* **51**: 26–42.

Braun, D. R., Plummer, T., Ferraro, J. V, Ditchfield, P. and Bishop, L. C. (2009b). Raw Material Quality and Oldowan Hominin Toolstone Preferences: Evidence from Kanjera South, Kenya. *Journal of Archaeological Science* **36**: 1605–1614.

Braun, D. R., Tactikos, J. C., Ferraro, J. V., Arnow, S. L. and Harris, J. W. K. (2008). Oldowan Reduction Sequences: Methodological Considerations. *Journal of Archaeological Science* **35**: 2153–2163.

Bretzke, K. and Conard, N. J. (2012). Evaluating Morphological Variability in Lithic Assemblages using 3D models of Stone Artifacts. *Journal of Archaeological Science* **39**: 3741–3749.

Brink, J. (1978). The Role of Abrasives in the Formation of Lithic Use-wear. *Journal of Archaeological Science* **5**: 363–371.

Briuer, F. L. (1976). New Clues to Stone Tool Function : Plant and Animal Residues. *American Antiquity* **41**: 478–484.

Brooks, A. S., Nevell, L., Yellen, J. E. and Hartman, G. (2006). Projectile Technologies of the African MSA. In E. Hovers and S. L. Kuhn (eds.), *Transitions Before the Transition: Evolution and Stability in the Middle Paleolithic and Middle Stone Age*, Springer, New York, pp.233–255.

Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments*, University of California Press, Berkeley.

Campbell, D. T. and Stanley, J. C. (1963). *Experimental and Quasi-Experimental Designs for Research*, Rand McNally, Chicago.

Carr, P. J. and Bradbury, A. P. (2001). Flake Debris Analysis, Levels of Production, and the Organization of Technology. In W. Andrefsky (ed.), *Lithic Debitage: Context, Form, Meaning*, University of Utah Press, Salt Lake City, pp.126–146.

Carr, P. J. and Bradbury, A. P. (2011). Learning from Lithics: A Perspective on the Foundation and Future of the Organization of Technology. *PaleoAnthropology* **2011**: 9–10.

Chase, P. G., Debénath, A., Dibble, H. L. and McPherron, S. P. (2009). *The Cave of Fontéchevade: Recent Excavations and Their Paleoanthropological Implications*, Cambridge University Press, Cambridge.

Chatters, J. C. (1987). Hunter-Gatherer Adaptations and Assemblage Structure. *Journal of Anthropological Archaeology* **6**: 336–375.

Cheshier, J. and Kelly, R. L. (2006). Projectile Point Shape and Durability: The Effect of Thickness:Length. *American Antiquity* **71**: 353–363.

Chiotti, L., Olszewski, D. I., Dibble, H. L., McPherron, S. P., Schurmans, U. A. and Smith, J. R. (2007). Paleolithic Abydos: Reconstructing Individual Behaviors across the High Desert Landscape. In Z. Hawass and J. Richards (eds.), *The Archaeology and Art of Ancient Egypt: Essays in Honor of David B. O'Connor*, The American University in Cairo Press, Cairo, pp.169–183.

Christenson, A. L. (1986). Projectile Point Size and Projectile Aerodynamics: An Exploratory Study. *Plains Anthropologist* **31**: 109–128.

Clarke, D. (1968). *Analytical Archaeology*, Methuen Publishing Ltd, London.

Clarke, D. (1973). Archaeology: The Loss of Innocence. *Antiquity* **47**: 6–18.

Clarkson, C. J. (2002). An Index of Invasiveness for the Measurement of Unifacial and Bifacial Retouch: A Theoretical, Experimental and Archaeological Verification. *Journal of Archaeological Science* **29**: 65–75.

Clarkson, C. J. (2010). Regional Diversity within the Core Technology of the Howiesons Poort Techno-Complex. In S. J. Lycett and P. R. Chauhan (eds.), *New Perspectives on Old Stones: Analytical Approaches in Paleolithic Technologies*, Springer, New York, pp.43–59.

Clarkson, C. J., Jones, S. and Harris, C. (2012). Continuity and Change in the Lithic Industries of the Jurreru Valley, India, Before and After the Toba Eruption. *Quaternary International* **258**: 165–179.

Close, A. E. (2000). Reconstructing Movement in Prehistory. *Journal of Archaeological Method and Theory* **7**: 49–77.

Cole, J. (1973). *Archaeology by Experiment*, Hutchinson, London.

Cole, J. (1979). *Experimental Archaeology*, Academic Press, London.

Collard, M. and Shennan, S. J. (2008). Patterns, Processes, and Parsimony: Studying Cultural Evolution with Analytical Techniques from Evolutionary Biology. In M. T. Stark B. J. Bowser and L. Horne (eds.), *Cultural Transmission and Material Culture: Breaking down Boundaries*, University of Arizona Press, Tuscon, pp.17–33.

Comis, L. (2010). Experimental Archaeology: Methodology and New Perspectives in Archaeological Open Air Museums. *Journal for (Re)construction and Experiment in Archaeology* **7**: 9–12.

Commins, S. and Linscott, R. N. (1947). *Man and the Universe: The Philosophers of Science*, Random House, New York.

Cook, T. D. and Campbell, D. T. (1979). *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Houghton Mifflin Company, Boston.

Cotterell, B. and Kamminga, J. (1987). The Formation of Flakes. *American Antiquity* **52**: 675–708.

Cotterell, B., Kamminga, J. and Dickson, F. P. (1985). The Essential Mechanics of Conchoidal Flaking. *International Journal of Fracture* **29**: 205–221.

Crabtree, D. E. and Davis, E. L. (1968). Experimental Manufacture of Wooden Implements with Tools of Flaked Stone. *Science* **159**: 426–8.

Crew, H. L. (1976). The Mousterian Site of Rosh Ein Mor. In A. E. Marks (ed.), *Prehistory and Paleoenvironments in the Central Negev, Israel. Volume I: The Avdat/Aqev Area, Part 1*, Southern Methodist University Press, Dallas, pp.75–112.

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.

Davidson, I. (1991). The Archaeology of Language Origins – A Review. *Antiquity* **65**: 39–48.

Davidson, I. (2002). The "Finished Artefact Fallacy": Acheulean Handaxes and Language Origins. In A. Wray (ed.), *The Transition to Language*, Oxford University Press, Oxford, pp.180–203.

Davidson, I. and Noble, W. (1993). Tools and Language in Human Evolution. In K. R. Gibson and T. Ingold (eds.), *Tools, Language and Cognition in Human Evolution*, Cambridge University Press, Cambridge, pp.363–388.

Debénath, A. and Dibble, H. L. (1994). *Handbook of Paleolithic Typology Vol. 1: The Lower and Middle Paleolithic of Europe*, University Museum Press, Philadelphia.

Dediu, D. and Levinson, S. C. (2013). On the Antiquity of Language: The Reinterpretation of Neandertal Linguistic Capacities and its Consequences. *Frontiers in Psychology* **4**: 397.

Delagnes, A. (2010). Du Rôle Structurant de la Mobilité dans les Systèmes Techniques du Paléolithique Moyen. In N. J. Conard and A. Delagnes (eds.), *Settlement Dynamics of the Middle Paleolithic and Middle Stone Age*, Kerns Verlag, Tubingen, pp.373–396.

Delagnes, A., Jaubert, J. and Meignen, L. (2007). Les Technocomplexes du Paléolithique Moyen en Europe Occidentale dans leur Cadre Diachronique et Géographique. In B. Vandermeersch and B. Maureille (eds.), *Les Neandertaliens. Biologie et Cultures*, Comité des Travaux Historiques et Scientifiques, Paris, pp.213–229.

Delagnes, A. and Meignen, L. (2006). Systems During the Middle Paleolithic in France Are There Any Chronological Trends? In E. Hovers and S. L. Kuhn (eds.), *Transitions Before the Transition: Evolution and Stability in the Middle Paleolithic and Middle Stone Age*, Springer, New York, pp.85–107.

Delagnes, A. and Rendu, W. (2011). Shifts in Neandertal Mobility, Technology and Subsistence Strategies in Western France. *Journal of Archaeological Science* **38**: 1771–1783.

Dibble, H. L. (1984). Interpreting Typological Variation of Middle Paleolithic Scrapers: Function, Style, or Sequence of Reduction. *Journal of Field Archaeology* **11**: 431–436.

Dibble, H. L. (1987). The Interpretation of Middle Paleolithic Scraper Morphology. *American Antiquity* **52**: 109–117.

Dibble, H. L. (1989). The Implications of Stone Tool Types for the Presence of Language during the Lower and Middle Paleolithic. In P. Mellars and C. B. Stringer (eds.), *The*

Human Revolution: Behavioural and Biological Perspectives on the Origins of Modern Humans, Edinburgh University Press, Edinburgh, pp.415–432.

Dibble, H. L. (1995a). Biache Saint-Vaast, Level IIA: A Comparison of Analytical Approaches. In H. L. Dibble and O. Bar-Yosef (eds.), *The Definition and Interpretation of Levallois Technology*, Prehistory Press, Madison, pp.93–116.

Dibble, H. L. (1995b). Middle Paleolithic Scraper Reduction: Background, Clarification, and Review of the Evidence to Date. *Journal of Archaeological Method and Theory* **2**: 299–368.

Dibble, H. L. (1995c). Raw Material Availability, Intensity of Utilizationm and Middle Paleolithic Assemblage Variability. *The Middle Paleolithic Site of Combe-Capelle Bas, France*, University Museum Press, Philadelphia, pp.289–316.

Dibble, H. L. (1997). Platform Variability and Flake Morphology: A Comparison of Experimental and Archaeological Data and Implications for Interpreting Prehistoric Lithic Technological. *Lithic Technology* **22**: 150–170.

Dibble, H. L., Aldeias, V., Alvarez-Fernandez, E., Blackwell, B. A. B., Hallett-Desguez, E., Jacobs, Z., Goldberg, P., Lin, S. C., Morala, A., Meyer, M. C., Olszewski, D. I., Reed, K., Reed, D., Richter, D., Roberts, R. G., Sandgathe, D. M., Schurmans, U. A., Skinner, A. R., Steele, T. E. and El-Hajraoui, M. (2012). New Excavations at the Site of Contrebandiers Cave, Morocco. *PaleoAnthropology* 145–201.

Dibble, H. L., Aldeias, V., Jacobs, Z., Olszewski, D. I., Rezek, Z., Lin, S. C., Alvarez-Fernández, E., Barshay-Szmidt, C. C., Hallett-Desguez, E., Reed, D., Reed, K., Richter, D., Steele, T. E., Skinner, A., Blackwell, B., Doronicheva, E. and El-Hajraoui, M. (2013). On the Industrial Attributions of the Aterian and Mousterian of the Maghreb. *Journal of Human Evolution* **2012**:

Dibble, H. L. and Bar-Yosef, O. eds. (1995). *The Definition and Interpretation of Levallois Technology*, Prehistory Press, Madison.

Dibble, H. L., Berna, F., Goldberg, P., McPherron, S. P., Mentzer, S., Niven, L., Richter, D., Sandgathe, D. M., Thery-Parisot, I. and Turq, A. (2009). A Preliminary Report on Pech de l'Aze IV, Layer 8 (Middle Paleolithic, France). *PaleoAnthropology* **2009**: 182–219.

Dibble, H. L. and Lenoir, M. eds. (1995). *The Middle Paleolithic site of Combe-Capelle Bas (France)*, University Museum Press, Philadelphia.

Dibble, H. L. and McPherron, S. P. (2006). The Missing Mousterian. *Current Anthropology* **47**: 777–803.

Dibble, H. L. and McPherron, S. P. (2007). Truncated-Faceted Pieces: Hafting Modification, Retouch, or Cores? In S. P. McPherron (ed.), *Tools versus Cores Alternative Approaches to Stone Tool Analysis*, Cambridge Scholars Publishing, Newcastle, pp.75–90.

Dibble, H. L., McPherron, S. P., Chase, P. G., Farrand, W. R. and Debénath, A. (2006). Taphonomy and the Concept of Paleolithic Cultures: The Case of the Tayacian from Fontéchevade. *PaleoAnthropology* 1–21.

Dibble, H. L. and Pelcin, A. (1995). The Effect of Hammer Mass and Velocity of Flake Mass. *Journal of Archaeological Science* **22**: 429–439.

Dibble, H. L. and Rezek, Z. (2009). Introducing a New Experimental Design for Controlled Studies of Flake Formation: Results for Exterior Platform Angle, Platform Depth, Angle of Blow, Velocity, and Force. *Journal of Archaeological Science* **36**: 1945–1954.

Dibble, H. L. and Rolland, N. (1992). On Assemblage Variability in the Middle Paleolithic of Western Europe. In H. L. Dibble and P. Mellars (eds.), *The Middle Paleolithic: Adaptation, Behavior, and Variability*, The University of Pennsylvania Museum of Archaeology and Anthropology, Philadelphia, pp.1–28.

Dibble, H. L., Schurmans, U. A., Radu, P. and Mclaughlin, M. V (2005). The Measurement and Interpretation of Cortex in Lithic Assemblages. *American Antiquity* **70**: 545–560.

Dibble, H. L. and Whittaker, J. C. (1981). New Experimental Evidence on the Relation between Percussion Flaking and Flake Variation. *Journal of Archaeological Science* **8**: 283–296.

Dobres, M.-A. (2000). *Technology and Social Agency: Outlining a Practice Framework for Archaeology*, Blackwell, Oxford.

Dobres, M.-A. and Hoffman, C. R. (1994). Social Agency and the Dynamics of Prehistoric Technology. *Journal of Archaeological Method and Theory* **1**: 211–258.

Domínguez-Rodrigo, M. (2008). Conceptual Premises in Experimental Design and their Bearing on the Use of Analogy: An Example from Experiments on Cut Marks. *World Archaeology* **40**: 67–82.

Douglass, M. J. (2010). *The Archaeological Potential of Informal Lithic Technologies: A Case Study of Assemblage Variability in Western New South Wales, Australia*, unpublished Ph.D. thesis, University of Auckland.

Douglass, M. J. and Holdaway, S. J. (2011). Quantifying Stone Raw Material Size Distributions: Investigating Cortex Proportions in Lithic Assemblages from Western New South Wales. In J. Specht and R. Torrence (eds.), *Changing Perspectives in Australian Archaeology, Part IV.*, Australiam Museum, Sydney, pp.45–57.

Douglass, M. J., Holdaway, S. J., Fanning, P. C. and Shiner, J. I. (2008). An Assessment and Archaeological Application of Cortex Measurement in Lithic Assemblages. *American Antiquity* **73**: 513–526.

Dunnell, R. C. (1980). Evolutionary Theory and Archaeology. *Advances in Archaeological Method and Theory* **3**: 35–99.

Dunnell, R. C. (1982). Science, Social Science, and Common Sense: The Agonizing Dilemma of Modern Archaeology. *Journal of Anthropological Research* **38**: 1–25.

Dunnell, R. C. (1989). Aspects of the Application of Evolutionary Theory in Archaeology. In C. C. Lamberg-Karlovsky (ed.), *Archaeological Thought in America*, Cambridge University Press, Cambridge, pp.35–99.

Dunnell, R. C. (1992). The Notion Site. In J. Rossignol and Wandsni (eds.), *Space, Time, and Archaeological Landscapes*, Springer, New York, pp.21–41.

Elston, R. G. (1990). A Cost-Benefit Model of Lithic Assemblage Variability. In R. G. Elston and E. E. Budy (eds.), *The Archaeology of James Creek Shelter*, University of Utah Press, Salt Lake City, pp.153–163.

Eren, M. I. and Bradley, B. A. (2009). Experimental Evaluation on the Levallois "Core Shape Maintenance" Hypothesis. *Lithic Technology* **34**: 119–125.

Eren, M. I., Greenspan, A. and Sampson, C. G. (2008). Are Upper Paleolithic Blade Cores More Productive than Middle Paleolithic Discoidal Cores? A Replication Experiment. *Journal of Human Evolution* **55**: 952–61.

Eren, M. I. and Lycett, S. J. (2012). Why Levallois? A Morphometric Comparison of Experimental "Preferential" Levallois Flakes versus Debitage Flakes. *PloS one* **7**: e29273.

Eren, M. I., Lycett, S. J., Roos, C. I. and Sampson, C. G. (2011). Toolstone Constraints on Knapping Skill: Levallois Reduction with Two Different Raw Materials. *Journal of Archaeological Science* **38**: 2731–2739.

Evens, J. (1872). *The Ancient Stone Implements, Weapons, and Ornaments of Great Britain*, D. Appleton, New York.

Faivre, J., Discamps, E., Gravina, B., Turq, A., Guadelli, J.-L. and Lenoir, M. (forthcoming). The Contribution of Lithic Production Systems to the Interpretation of Mousterian Industrial Variability in South-Western France: The example of Combe-Grenal (Dordogne, France). *Quaternary International*.

Faulkner, A. (1972). *Mechanical Principles of Flint Working*, unpublished Ph.D. dissertation, Washington State University.

Feblot-Augustins, J. (1993). Mobility Strategies in the Late Middle Paleolithic of Central Europe and Western Europe: Elements of Stability and Variability. *Journal of Anthropological Archaeology* **12**: 211–265.

Feblot-Augustins, J. (1999). Raw Material Transport Patterns and Settlement Systems in the European Lower and Middle Palaeolithic: Continuity, Change and Variability. In W. Roebroeks and C. Gamble (eds.), *The Middle Paleolithic Occupation of Europe*, Leiden University Press, Leiden, pp.193–214.

Fernandes, P., Raynal, J.-P. and Moncel, M.-H. (2008). Middle Palaeolithic Raw Material Gathering Territories and Human Mobility in the Southern Massif Central, France: First Results from a Petro-Archaeological Study on Flint. *Journal of Archaeological Science* **35**: 2357–2370.

Ferring, R. (1977). The Late Upper Paleolithic Site of Ein Aqev East. In A. E. Marks (ed.), *Prehistory and Paleoenvironments in the Central Negev, Israel. Volume II: The Awiat/Aqev Area, Part 2 and the Har Harif*, Southern Methodist University Press, Dallas, pp.81–118.

Fischer, A., Hansen, P. V. and Rasmussen, P. (1984). Macro and Micro Wear Traces on Lithic Projectile Points. *Journal of Danish Archaeology* **3**: 19–46.

Flannery, K. V. (1973). Archaeology with A Capital "S." In C. Gans (ed.), *Research and Theory in Current Archaeology*, Wiley, New York, pp.47–53.

Flenniken, J. J. (1978). Reevaluation of the Lindenmeier Folsom: A Replication Experiment in Lithic Technology. *American Antiquity* **43**: 473–480.

Flenniken, J. J. (1984). The Past, Present, and Future of Flintknapping: An Anthropological Perspective. *Annual Review of Anthropology* **13**: 187–203.

Flenniken, J. J. (1987). The Paleolithic Dyuktai Pressure Blade Technique of Siberia. *Arctic Anthropology* **24**: 117–132.

Flenniken, J. J. and Raymond, A. W. (1986). Morphological Projectile Point Typology: Replication Experimentation and Technological Analysis. *American Antiquity* **51**: 603–614.

Fogelin, L. (2007). Inference to the Best Explanation: A Common and Effective From of Archaeological Reasoning. *American Antiquity* **72**: 603–625.

Foley, R. and Lahr, M. M. (2003). On Stony Ground: Lithic Technology, Human Evolution, and the Emergence of Culture. *Evolutionary Anthropology: Issues, News, and Reviews* **12**: 109–122.

Ford, J. A. (1954a). Comment on A.C. Spaulding's "Statistical Techniques for the Discovery of Artifact Types." *American Antiquity* **19**: 390–391.

Ford, J. A. (1954b). On the Concept of Types: The Type Concept Revisited. *American Anthropologist* **56**: 42–53.

Ford, J. A. (1954c). Spaulding's Review of Ford. *American Anthropologist* **56**: 109–114.

Forestier, H. (1993). Le Clactonien: Mise en Application d'une Nouvelle Méthode de Débitage s'inscrivant dans la Variabilité des Systèmes de Production Lithique du Paléolithique Ancien. *Paléo* **5**: 53–82.

Frison, G. C. (1968). A Functional Analysis of Certain Chipped Stone Tools. *American Antiquity* **33**: 149–155.

Frison, G. C. and Raymond, A. (1980). *Folsom Tools and Technology at the Hanson Site*, University of New Mexico Press, Albuquerque.

Fritz, J. M. and Plog, F. T. (1970). The Nature of Archaeological Explanation. *American Antiquity* **35**: 405–412.

Gauch, H. G. (2003). *Scientific Method in Practice*, Cambridge University Press, Cambridge.

Geneste, J.-M. (1985). *Analyse Lithique d'Industries Mousteriennes du Perigord: Une Approache Technologique du Comportement des Groupes Humaines au Paleolithique Moyen*, unpublished Ph.D. thesis, Université de Bordeaux.

Geneste, J.-M. (1991). Systèmes Techniques de Production Lithique: Variations Technoéconomiques dans les Processus de Réalisation des Outillages Paléolithiques. *Techniques & Culture* **17-18**: 1–35.

Geribàs, N., Mosquera, M. and Vergès, J. M. (2010). What Novice Knappers have to Learn to Become Expert Stone Toolmakers. *Journal of Archaeological Science* **37**: 2857–2870.

Gibbon, G. E. (1989). *Explanation in Archaeology*, Blackwell, New York.

Gifford-Gonzalez, D. (1989). Modem Analogues: Developing an Interpretive Framework. In R. Bonnichsen and M. Sorg (eds.), *Bone Modification*, Center for the Study of the First Americans, University of Maine, Orono, pp.43–52.

Gifford-Gonzalez, D. (1991). Bones are Not Enough: Analogues, Knowledge, and Interpretive Strategies in Zooarchaeology. *Journal of Anthropological Archaeology* **10**: 215–254.

Goldberg, P., Dibble, H. L., Berna, F., Sandgathe, D. M., McPherron, S. P. and Turq, A. (2012). New Evidence on Neandertal Use of Fire: Examples from Roc de Marsal and Pech de l'Azé IV. *Quaternary International* **247**: 325–340.

Gould, R. A. (1980). *Living Archaeology*, Cambridge University Press, Cambridge.

Gould, R. A. and Watson, P. J. (1982). A Dialogue on the Meaning and Use of Analogy in Ethnoarchaeological Reasoning. *Journal of Anthropological Archaeology* **1**: 355–381.

Gould, S. J. (1965). Is Uniformitarianism Necessary? *American Journal of Science* **236**: 223–8.

Grayson, D. K. and Cole, S. C. (1998). Stone Tool Assemblage Richness during the Middle and Early Upper Palaeolithic in France. *Journal of Archaeological Science* **25**: 927–938.

Guérin, G., Discamps, E., Lahaye, C., Mercier, N., Guibert, P., Turq, A., Dibble, H. L., McPherron, S. P., Sandgathe, D., Goldberg, P., Jain, M., Thomsen, K., Patou-Mathis, M., Castel, J.-C. and Soulier, M.-C. (2012). Multi-Method (TL and OSL), Multi-Material (quartz and flint) Dating of the Mousterian site of Roc de Marsal (Dordogne, France): Correlating Neanderthal Occupations with the Climatic Variability of MIS 5–3. *Journal of Archaeological Science* **39**: 3071–3084.

Guibert, P., Lahaye, C. and Bechtel, F. (2009). The Importance of U-Series Disequilibrium of Sediments in Luminescence Dating: A Case Study at the Roc de Marsal Cave (Dordogne, France). *Radiation Measurements* **44**: 223–231.

Haidel, M. N. (2007). Archaeology. In W. Henke and I. Tattersall (eds.), *Handbook of Paleoanthropology*, Springer, New York, pp.261–287.

Harris, J. W. K. and Isaac, G. L. (1997). Sites in the Upper KBS, Okote, and Chari members. In G. L. Isaac (ed.), *Koobi Fora Research Project*, Clarendon Press, Oxford, pp.115–220.

Haslam, M., Hernandez-Aguilar, A., Ling, V., Carvalho, S., de la Torre, I., DeStefano, A., Du, A., Hardy, B., Harris, J., Marchant, L., Matsuzawa, T., McGrew, W., Mercader, J., Mora, R., Petraglia, M., Roche, H., Visalberghi, E. and Warren, R. (2009). Primate Archaeology. *Nature* **460**: 339–44.

Hawkes, C. (1954). Archaeological Theory and Method: Some Suggestions from the Old World. *American Anthropologist* **56**: 155–168.

Hayden, B. ed. (1979a). *Lithic Use-Wear Analysis*, Academic Press, New York.

Hayden, B. (1979b). *Paleolithic Reflections*, Australian Institute of Aboriginal Studies, Canberra.

Hempel, C. G. (1965). *Aspects of Scientific Explanation, and Other Essays in the Philosophy of Science*, The Free Press, New York.

Henke, W. and Tattersall, I. eds. (2007). *Handbook of Paleoanthropology*, Springer-Verlag, Berlin.

Henry, D. O., Haynes, C. V. and Bradley, B. A. (1976). Quantitative Variations in Flaked Stone Debitage. *Plains Anthropologist* **21**: 57–61.

Henshilwood, C. S. and Dubreuil, B. (2011). The Still Bay and Howiesons Poort, 77–59 ka. *Current Anthropology* **52**: 361–400.

Hiscock, P. (2001). Sample Size and Composition of Artefact Assemblages. *Australian Aboriginal Studies* **1**: 48–62.

Hiscock, P. (2004). Slippery and Billy: Intention, Selection and Equifinality in Lithic Artefacts. *Cambridge Archaeological Journal* **14**: 71–77.

Hiscock, P., Turq, A., Faivre, J. and Bourguignon, L. (2009). Quina Procurement and Tool Production. In B. Adams and B. S. Blades (eds.), *Lithic Materials and Paleolithic Societies*, Wiley Blackwell, New York.

Hodder, I. (1999). *The Archaeological Process: An Introduction*, Blackwell, Oxford.

Hodder, I. (2012). Introduction: Contemporary Theoretical Debate. In I. Hodder (ed.), *Archaeological Theory Today*, Polity Press, Cambridge, pp.1–14.

Holdaway, S. J. and Allen, H. (2013). Placing Ideas in the Land: Practical and Ritual Training Among the Australian Aborigines. In W. Wendrich (ed.), *Archaeology and Apprenticeship*, University of Arizona Press, Tuscon, pp.79–98.

Holdaway, S. J. and Douglass, M. J. (2012). A Twenty-First Century Archaeology of Stone Artifacts. *Journal of Archaeological Method and Theory* **19**: 101–131.

Holdaway, S. J., Douglass, M. J. and Fanning, P. C. (2012). Landscape Scale and Human Mobility: Geoarchaeological Evidence from Rutherfords Creek, New South Wales, Australia. In S. J. Kluiving and E. B. Guttmann-Bond (eds.), *Landscape Archaeological between Art and Science: From a Multi- to an Interdisciplinary Approach*, Amsterdam University Press, Amsterdam, pp.279–294.

Holdaway, S. J., Douglass, M. J. and Fanning, P. C. (2013). A New Ecological Framework for Understanding Human-Environment Interaction in Arid Australia. In D. Frankel J. M. Webb and S. Lawrence (eds.), *Archaeology in Environment and Technology: Intersections and Transformations*, Routledge, pp.51–68.

Holdaway, S. J., Shiner, J. I., Fanning, P. C. and Douglass, M. J. (2008). Assemblage Formation as a Result of Raw Material Acquisition in Western New South Wales, Australia. *Lithic Technology* **33**: 73–85.

Holdaway, S. J. and Wandsnider, L. eds. (2008). *Time in Archaeology: Time Perspectivism Revisited*, University of Utah Press, Salt Lake City.

Holdaway, S. J., Wendrich, W. and Phillipps, R. (2010). Identifying Low-Level Food Producers: Detecting Mobility from Lithics. *Antiquity* **84**: 185–194.

Holmes, W. H. (1894). Natural History of Flaked Stone Implements. *Memoirs of the International Congress of Anthropology*, Chicago, pp.120–139.

Hughes, R. E. (1998). On Realiability, Validity, and Scale in Obsidian Sourcing Research. In A. F. Ramenofsky and A. Steffen (eds.), *Unit Issues in Archaeology: Measuring Time, Space, and Material*, University of Utah Press, Salt Lake City, pp.103–116.

Ingbar, E. E., Larson, M. L. and Bradley, B. A. (1989). A Non-Typological Approach to Debitage Analysis. In D. S. Amick and R. P. Mauldin (eds.), *Experiments in Lithic Technology*, BAR International Series, Oxford, pp.67–99.

Ingersoll, D. and MacDonald, W. (1977). Introduction. In D. Ingersoll, J. E. Yellen, and W. K. Macdonald (eds.), *Experimental Archaeology*, Columbia University Press, New York, pp.xi–xvii.

Inizan, M.-L., Reduron, M., Roche, H. and Tixier, J. (1995). *Techologie de la Pierre Taillée, Tome 4*, CREP, Meudon.

Iovita, R. (2009). Ontogenetic Scaling and Lithic Systematics: Method and Application. *Journal of Archaeological Science* **36**: 1447–1457.

Iovita, R. (2011). Shape Variation in Aterian Tanged Tools and the Origins of Projectile Technology: A Morphometric Perspective on Stone Tool Function. *PloS one* **6**: e29029.

Iovita, R. and McPherron, S. P. (2011). The Handaxe Reloaded: A Morphometric Reassessment of Acheulian and Middle Paleolithic Handaxes. *Journal of Human Evolution* **61**: 61–74.

Iovita, R., Schönekeß, H., Gaudzinski-Windheuser, S. and Jäger, F. (2014). Projectile Impact Fractures and Launching Mechanisms: Results of a Controlled Ballistic Experiment Using Replica Levallois Points. *Journal of Archaeological Science* **48**: 73–83.

Isaac, G. L. and Harris, J. W. K. (1997). The Stone Artifact Assemblages: A Comparative Study. In G. L. Isaac (ed.), *Koobi Fora Research Project*, Clarendon Press, Oxford, pp.262–362.

Jennings, T. a., Pevny, C. D. and Dickens, W. a. (2010). A Biface and Blade Core Efficiency Experiment: Implications for Early Paleoindian Technological Organization. *Journal of Archaeological Science* **37**: 2155–2164.

Jeske, R. J. (1989). Economies in Raw Material Use by Prehistoric Hunter-Gatherers. In R. Torrence (ed.), *Time, Energy and Stone Tools*, Cambridge University Press, Cambridge, pp.34–45.

Jia, P. W., Doelman, T., Chen, C., Zhao, H., Lin, S., Torrence, R. and Glascock, M. D. (2010). Moving Sources: A Preliminary Study of Volcanic Glass Artifact Distributions in Northeast China Using PXRF. *Journal of Archaeological Science* **37**: 1670–1677.

Johnson, L. L. (1978). A History of Flint-Knapping Experimentation, 1838-1976. *Current Anthropology* **19**: 337–372.

Kamminga, J. (1980). A Functional Investigation of Australian Microliths. *The Artefact* **5**: 1–18.

Kelly, R. L. (1988). The Three Sides of a Biface. *American Antiquity* **53**: 717–734.

Kelly, R. L. and Todd, L. C. (1988). Coming into the Country : Early Paleoindian Hunting and Mobility. *American Antiquity* **53**: 231–244.

Key, A. J. M. and Lycett, S. J. (2011). Technology Based Evolution? A Biometric Test of the Effects of Handsize Versus Tool Form on Efficiency in an Experimental Cutting Task. *Journal of Archaeological Science* **38**: 1663–1670.

Kirk, R. E. (2009). Experimental Design. In R. E. Millsap and A. Maydeu-Olivares (eds.), *The SAGE Handbook of Quantitative Methods in Psychology*, SAGE Publications, Thousand Oaks, pp.23–45.

Kirk, R. E. (2012). *Experimental Design: Procedures for the Behavioral Sciences, 4th ed.*, SAGE Publications, Thousand Oaks.

Kobayashi, H. (1975). The Experimental Study of Bipolar Flakes. In E. H. Swanson (ed.), *Lithic Technology: Making and Using Stone Tools*, Mouton Publishers, The Hague, pp.115–128.

Kosso, P. (2011). *A Summary of Scientific Method*, Springer, New York.

Kuhn, S. L. (1990). A Geometric Index of Reduction for Unifacial Stone Tools. *Journal of Archaeological Science* **17**: 583–593.

Kuhn, S. L. (1991). " Unpacking " Reduction : Lithic Raw Material Economy in the Mousterian of West-Central Italy. *Journal of Anthropological Archaeology* **10**: 76–106.

Kuhn, S. L. (1992). On Planning and Curated Technologies in the Middle Paleolithic. *Journal of Anthropological Research* **48**: 185–214.

Kuhn, S. L. (1994). A Formal Approach to the Design and Assembly of Mobile Toolkits. *American Antiquity* **59**: 426–442.

Kuhn, S. L. (1995). *Mousterian Lithic Technology: An Ecological Perspective*, Princeton University Press, Princeton.

Kuhn, S. L. (2004a). Middle Paleolithic Assemblage Formation at Riparo Mochi. In A. L. Johnson (ed.), *Processual Archaeology: Exploring Analytical Strategies, Frames of Reference, and Culture Process*, Praeger, Westport, pp.31–60.

Kuhn, S. L. (2004b). Upper Paleolithic Raw Material Economies at Üçağızlı cave, Turkey. *Journal of Anthropological Archaeology* **23**: 431–448.

Kuhn, S. L. and Zwyns, N. (2014). Rethinking the Initial Upper Paleolithic. *Quaternary International* **347**: 29–38.

Lalueza-Fox, C. and Gilbert, M. T. P. (2011). Paleogenomics of Archaic Hominins. *Current Biology* **21**: R1002–R1009.

Lalueza-Fox, C., Rosas, A., de la Rasilla, M., Gilbert, M. T. P. and Willerslev, E. (2011). Reply to Vigilant and Langergraber: Patrilocality in Neandertals is Still the Most Plausible Explanation. *Proceedings of the National Academy of Sciences of the United States of America* **108**: E88–E88.

Lerner, H. (2013). Experiments and their Application to Lithic Archaeology: An Experimental Essay. In S. Chrisomalis and A. Costopoulos (eds.), *Human Expeditions: Inspired by Bruce Trigger*, university of Toronto Press, Toronto, pp.73–89.

Leroi-Gourhan, A. (1964). *Le Geste et la Parole 1. Techniques et Langage*, Aldine, Paris.

Li, Y., Hou, Y. and Boëda, É. (2009). Mode of Débitage and Technical Cognition of Hominids at the Guanyindong site. *Chinese Science Bulletin* **54**: 3864–3871.

Lin, S. C. H., Douglass, M. J., Holdaway, S. J. and Floyd, B. (2010). The Application of 3D Laser Scanning Technology to the Assessment of Ordinal and Mechanical Cortex Quantification in Lithic Analysis. *Journal of Archaeological Science* **37**: 694–702.

Lin, S. C., McPherron, S. P. and Dibble, H. L. (forthcoming). Establishing Statistical Confidence in Cortex Ratios between Lithic Assemblages: A Case Study of the Middle Paleolithic of southwestern France. *Journal of Archaeological Science*.

Lin, S. C., Rezek, Z., Braun, D. R. and Dibble, H. L. (2013). On the Utility and Economization of Unretouched Flakes: The Effects of Exterior Platform Angle and Platform Depth. *American Antiquity* **78**: 724–745.

Lipo, C. P., Dunnell, R. C., O'Brien, M. J., Harper, V. and Dudgeon, J. (2012). Beveled Projectile Points and Ballistics Technology. *American Antiquity* **77**: 774–788.

Lombard, M. (2004). Distribution Patterns of Organic Residues on Middle Stone Age Points from Sibudu Cave, Kwazulu-Natal, South Africa. *The South African Archaeological Bulletin* **59**: 37–44.

Lombard, M. (2005). Evidence of Hunting and Hafting during the Middle Stone Age at Sibidu Cave, KwaZulu-Natal, South Africa: A Multianalytical Approach. *Journal of Human Evolution* **48**: 279–300.

Lombard, M. (2006). Direct Evidence for the Use of Ochre in the Hafting Technology of Middle Stone Age tools from Sibudu Cave. *Southern African Humanities* **18**: 57–67.

Lombard, M. (2007). The Gripping Nature of Ochre: The Association of Ochre with Howiesons Poort Adhesives and Later Stone Age Mastics from South Africa. *Journal of Human Evolution* **53**: 406–19.

Lombard, M. (2008). Finding Resolution for the Howiesons Poort through the Microscope: Micro-Residue Analysis of Segments from Sibudu Cave, South Africa. *Journal of Archaeological Science* **35**: 26–41.

Lombard, M. and Pargeter, J. (2008). Hunting with Howiesons Poort Segments: Pilot Experimental Study and the Functional Interpretation of Archaeological Tools. *Journal of Archaeological Science* **35**: 2523–2531.

Longacre, W. A. (1970). *Archaeology as Anthropology: A Case Study*, University of Arizona Press, Tuscon.

Lycett, S. J. (2007). Why is there a Lack of Mode 3 Levallois Technologies in East Asia? A Phylogenetic Test of the Movius–Schick hypothesis. *Journal of Anthropological Archaeology* **26**: 541–575.

Lycett, S. J. (2008). Acheulean Variation and Selection: Does Handaxe Symmetry Fit Neutral Expectations? *Journal of Archaeological Science* **35**: 2640–2648.

Lycett, S. J. (2009a). Are Victoria West Cores "Proto-Levallois"? A Phylogenetic Assessment. *Journal of Human Evolution* **56**: 175–191.

Lycett, S. J. (2009b). Understanding Ancient Hominin Dispersals Using Artefactual Data: A Phylogeographic Analysis of Acheulean Handaxes. *PloS one* **4**: e7404.

Lycett, S. J. (2010). Cultural Transmission, Genetic Models and Paleolithic Variability: Integrative Analytical Appraches. In S. J. Lycett and P. R. Chauhan (eds.), *New Perspectives on Old Stones: Analytical Approaches in Paleolithic Technologies*, Springer, New York, pp.207–234.

Lycett, S. J. (2011). "Most Beautiful and Most Wonderful": Those Endless Stone Tool Forms. *Journal of Evolutionary Psychology* **9**: 143–171.

Lycett, S. J. and Chauhan, P. R. eds. (2010). *New Perspectives on Old Stones: Analytical Approaches to Paleolithic Technologies*, Springer, New York.

Lycett, S. J. and von Cramon-Taubadel, N. (forthcoming). Toward a "Quantitative Genetic" Approach to Lithic Variation. *Journal of Archaeological Method and Theory*.

Lycett, S. J. and von Cramon-Taubadel, N. (2008). Acheulean Variability and Hominin Dispersals: A Model-Bound Approach. *Journal of Archaeological Science* **35**: 553–562.

Lycett, S. J. and von Cramon-Taubadel, N. (2013). A 3D Morphometric Analysis of Surface Geometry in Levallois Cores: Patterns of Stability and Variability Across Regions and their Implications. *Journal of Archaeological Science* **40**: 1508–1517.

Lycett, S. J., von Cramon-Taubadel, N. and Gowlett, J. A. J. (2010). A Comparative 3D Geometric Morphometric Analysis of Victoria West Cores: Implications for the Origins of Levallois Technology. *Journal of Archaeological Science* **37**: 1110–1117.

Lycett, S. J. and Eren, M. I. (2013). Levallois Lessons: The Challenge of Integrating Mathematical Models, Quantitative Experiments and the Archaeological Record. *World Archaeology* **45**: 519–538.

Lycett, S. J. and Norton, C. J. (2010). A Demographic Model for Palaeolithic Technological Evolution: The Case of East Asia and the Movius Line. *Quaternary International* **211**: 55–65.

Lyman, R. L. and O'Brien, M. J. (1998). The Goals of Evolutionary Archaeology: History and Explanation. *Current Anthropology* **39**: 615–652.

Mackay, A. (2008). A Method for Estimating Edge Length from Flake Dimensions: Use and Implications for Technological Change in the Southern African MSA. *Journal of Archaeological Science* **35**: 614–622.

Magnani, M., Rezek, Z., Lin, S. C., Chen, A. and Dibble, H. L. (2014). Flake Variation in Relation to the Application of Force. *Journal of Archaeological Science* **46**: 37–49.

Magne, M. and Pokotylo, D. (1981). A Pilot Study in Bifacial Lithic Reduction Sequences. *Lithic Technology* **10**: 34–47.

Marks, A. E. (1968). The Khormusanan Upper Pleistocene Industry in Sudanese Nubia. In F. Wendorf (ed.), *The Prehistory of Nubia*, For Burgwin Research Center and Southern Methodist University, Dallas, pp.315–391.

Marks, A. E. (1976). Ein Aqev: A Late Levantine Upper Paleolithic Site in the Nahal Aqev. In A. E. Marks (ed.), *Prehistory and Paleoenvironments in the Central Negev,*

Israel. Volume I: The Avcfat/Aqev Area, Part 1, Southern Methodist University Press, Dallas, pp.227–292.

Marks, A. E. (1977). The Upper Paleolithic Sites of Boker Tachtit and Boker: A Preliminary Report. In A. E. Marks (ed.), *Prehistory and Paleoenvironments in the Central Negev, Israel. Volume II: The Avdat/Aqev Area, Part 2 and the Har Harif*, Southern Methodist University Press, Dallas, pp.61–80.

Marwick, B. (2008). What Attributes are Important for the Measurement of Assemblage Reduction Intensity? Results from an Experimental Stone Artefact Assemblage with Relevance to the Hoabinhian of Mainland Southeast Asia. *Journal of Archaeological Science* **35**: 1189–1200.

Mathieu, J. R. (2002). Introduction. In J. R. Mathieu (ed.), *Experimental Archaeology, Replicating Past Objects, Behaviours and Processes*, Archaeopress, Oxford, pp.1–12.

Matsuzawa, T. (2001). Primate Foundations of Human Intelligence: A View of Tool Use in Nonhuman Primates and Fossil Hominids. In T. Matsuzawa (ed.), *Primate Origins of Human Cognition and Behavior*, Springer Japan, Tokyo, pp.3–25.

McPherron, S. P. (1999). Ovate and Pointed Handaxe Assemblages: Two Points Make a Line. *Préhistoire Européenne* **14**: 9–32.

McPherron, S. P. (2013). Perspectives on Stone Tools and Cognition in the Early Paleolithic Record. In C. M. Sanz J. Call and C. Boesch (eds.), *Tool Use in Animals: Cognition and Ecology*, Cambridge University Press, Cambridge, pp.286–309.

McPherron, S. P. and Dibble, H. L. (2000). The Lithic Assemblages of Pech de l'Azé IV (Dordogne, France). *Préhistoire Européenne* **15**: 9–43.

McPherron, S. P., Talamo, S., Goldberg, P., Niven, L., Sandgathe, D., Richards, M. P., Richter, D., Turq, A. and Dibble, H. L. (2012). Radiocarbon Dates for the Late Middle Palaeolithic at Pech de l'Azé IV, France. *Journal of Archaeological Science* **39**: 3436–3442.

Meignen, L., Delagnes, A. and Bourguignon, L. (2009). Patterns of Lithic Material Procurement and Transformation During the Middle Paleolithic in Western Europe. In B. Adams and B. S. Blades (eds.), *Lithic Materials and Paleolithic Societies*, Wiley Blackwell, New York, pp.15–24.

Meltzer, D. J., Leonardb, R. D. and Strattonb, S. K. (1992). The Relationship Between Sample Size and Diversity in Archaeological Assemblages. *Journal of Archaeological Science* **19**: 375–387.

Montet-White, A. and Holen, S. eds. (1991). *Raw Material Economics among Prehistoric Hunter-Gatherers*, University of Kansas Publications in Anthropology 19, Lawrence.

Moore, M. W. (2003). Flexibility of Stone Tool Manufacturing Methods on the Georgina River, Camooweal, Queensland. *Archaeology in Oceania* **38**: 23–36.

Moore, M. W. (2007). Lithic Design Space Modelling and Cognition in Homo floresiensis. In A. C. Schalley and D. Khlentzos (eds.), *Mental Stages, Volume 1: Evolution, Function, and Nature*, pp.11–34.

Moore, M. W. (2011). The Design Space of Stone Flaking: Implications for Cognitive Evolution. *World Archaeology* **43**: 702–715.

Moore, M. W. and Brumm, A. (2009). Homo floresiensis and the African Oldowan. In E. Hovers and D. R. Braun (eds.), *Interdisciplinary Approaches to the Oldowan: Vertebrate Paleobiology and Paleoanthropology*, Springer Netherlands, Dordrecht, pp.61–69.

Morrow, J. (1997). End Scraper Morphology and Use-Life: An Approach for Studying Paleoindian Lithic Technology and Mobility. *Lithic Technology* **22**: 70–85.

Morrow, T. A. (1996). Bigger is Better: Comments on Kuhn's Formal approach to Mobile Tool Kits. *American Antiquity* **61**: 581–590.

Mourre, V., Villa, P. and Henshilwood, C. S. (2010). Early Use of Pressure Flaking on Lithic Artifacts at Blombos Cave, South Africa. *Science* **330**: 659–62.

Munday, F. C. (1976). Intersite Variability in the Mousterian Occupation of the Avdat/Aqev Area. In A. E. Marks (ed.), *Prehistory and Paleoenvironments in the Central Negev, Israel. Volume II: The Avdat/Aqev Area, Part 2 and the Har Harif*, Southern Methodist University Press, Dallas, pp.113–140.

Nelson, M. C. (1991). The Study of Technological Organization. *Archaeological Method and Theory* **3**: 57–100.

Newcomer, M. H. (1971). Some Quantitative Experiments in Handaxe Manufacture. *World Archaeology* **3**: 85–94.

Newcomer, M. H. (1975). "Punch Technique" and Upper Paleolithic Blades. In E. H. Swanson (ed.), *Lithic Technology: Making and Using Stone Tools*, Mouton Publishers, The Hague, pp.97–104.

Niiniluoto, I. (1999). Defending Abduction. *Philosophy of Science* **66**: S436–451.

Niven, L. (2013). A Diachronic Evaluation of Neanderthal Cervid Exploitation and Site Use at Pech de l'Azé IV, France. In J. L. Clark and J. D. Speth (eds.), *Zooarchaeology and Modern Human Origins: Human Hunting Behavior during the Later Pleistocene*, Springer, New York, pp.151–161.

Niven, L., Steele, T. E., Rendu, W., Mallye, J.-B., McPherron, S. P., Soressi, M., Jaubert, J. and Hublin, J.-J. (2012). Neandertal Mobility and Large-Game Hunting: The Exploitation of Reindeer During the Quina Mousterian at Chez-Pinaud Jonzac (Charente-Maritime, France). *Journal of Human Evolution* **63**: 624–635.

Nonaka, T., Bril, B. and Rein, R. (2010). How Do Stone Knappers Predict and Control the Outcome of Flaking? Implications for Understanding Early Stone Tool Technology. *Journal of Human Evolution* **59**: 155–67.

O'Brien, M. J., Darwent, J. and Lyman, R. L. (2001). Cladistics is Useful for Reconstructing Archaeological Phylogenies: Palaeoindian Points from the Southeastern United States. *Journal of Archaeological Science* **28**: 1115–1136.

O'Brien, M. J. and Holland, T. D. (1990). Variation, Selection, and the Archaeological Record. *Archaeological Method and Theory* **2**: 31–79.

O'Brien, M. J. and Lyman, R. L. (2002). The Epistemological Nature of Archaeological Units. *Anthropological Theory* **2**: 37–56.

O'Brien, M. J., Lyman, R. L., Mesoudi, A. and VanPool, T. L. (2010). Cultural Traits as Units of Analysis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **365**: 3797–806.

Odell, G. H. (1996). Economizing Behavior and the Concept of "Curation." In G. H. Odell (ed.), *Stone Tools: Theoretical Insights into Human Prehistory*, Plenum Press, New York, pp.51–80.

Olszewski, D. I., Dibble, H. L., McPherron, S. P., Schurmans, U. A., Chiotti, L. and Smith, J. R. (2010). Nubian Complex Strategies in the Egyptian High Desert. *Journal of Human Evolution* **59**: 188–201.

Outram, A. K. (2008). Introduction to Experimental Archaeology. *World Archaeology* **40**: 1–6.

Pargeter, J. (2011). Assessing the Macrofracture Method for Identifying Stone Age Hunting Weaponry. *Journal of Archaeological Science* **38**: 2882–2888.

Parker, D. (2011). *The Complexity of Lithic Simplicity: Computer Simulation of Lithic Assemblage Formation in Western New South Wales, Australia*, unpublished Master dissertation, University of Auckland.

Parry, W. J. and Kelly, R. L. (1987). Expedient Core Technology and Sedentism. In J. K. Johnson and C. A. Morrow (eds.), *The Organization of Core Technology*, Westview Press, Boulder, pp.285–304.

Pearce, E., Stringer, C. and Dunbar, R. I. M. (2013). New Insights into Differences in Brain Organization between Nenaderthals and Anatomically Modern Humans. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **280**: 20130168.

Van Peer, P. (1992). *The Levallois Reduction Strategy*, Prehistory Press, Madison.

Pelcin, A. W. (1997a). The Effect of Core Surface Morphology on Flake Attributes: Evidence from a Controlled Experiment. *Journal of Archaeological Science* **24**: 749–756.

Pelcin, A. W. (1997b). The Formation of Flakes: The Role of Platform Thickness and Exterior Platform Angle in the Production of Flake Initiations and Terminations. *Journal of Archaeological Science* **24**: 1107–1113.

Pelegrin, J. (1990). Prehistoric Lithic Technology: Some Aspects of Research. *Archaeological Review from Cambridge* **9**: 116–125.

Pelegrin, J. (2009). Cognition and the Emergence of Language: A Contribution from Lithic Technology. In S. de Beaune F. L. Coolidge and T. Wynn (eds.), *Cognitive Archaeology and Human Evolution*, Cambridge University Press, Cambridge, pp.95–108.

Perlés, C. (1992). In Search of Lithic Strategies: A Cognitive Approach to Prehistoric Chipped Stone Assemblages. In J.-C. Gardin and C. S. Peebles (eds.), *Representations in Archaeology*, Indiana University Press, Bloomington, pp.223–247.

Perreault, C., Brantingham, P. J., Kuhn, S. L., Wurz, S. and Gao, X. (2013). Measuring the Complexity of Lithic Technology. *Current Anthropology* **45**: S397–S406.

Phillipps, R. S. (2012). *Documenting Socio-Economic Variability in the Egyptian Neolithic Through Stone Artefact Analysis*, unpublished Ph.D. thesis, University of Auckland.

Pigeot, N. (1990). Technical and Social Actors: Flint Knapping Specialist at Magdalenian Etoilles. *Archaeological Review from Cambridge* **9**: 126–141.

Potts, R. (1991). Why the Oldowan? Plio-Pleistocene Toolmaking and the Transport of Resources. *Journal of Anthropological Research* **47**: 153–176.

Powell, A., Shennan, S. J. and Thomas, M. G. (2009). Late Pleistocene Demography and the Appearance of Modern Human Behavior. *Science* **324**: 1298–1301.

Prasciunas, M. M. (2007). Bifacial Cores and Flake Production Efficiency: An Experimental Test of Technological Assumptions. *American Antiquity* **72**: 334–348.

R Core Team (2013). R: A Language and Environment for Statistical Computing.

Raab, L. M. and Goodyear, A. C. (1984). Middle-Range Theory in Archaeology : A Critical Review of Origins and Applications. *American Antiquity* **49**: 255–268.

Reher, C. A. (1991). Large Scale Lithic Quarries and Regional Transport Systems on the High Plains of Eastern Wyoming: Spanish Diggings Revisited. In A. Montet-White and S. Holen (eds.), *Raw Matrerial Economies among Prehistoric Hunter-Gaterers*, University of Kansas Publications in Anthropology 19, Lawrence, pp.251–284.

Rezek, Z., Lin, S. C. and Dibble, H. L. (forthcoming). The Role of Controlled Experiments in Understanding Variation in Flake Production. In D. I. Olszewski and A. P. Sullivan (eds.), *Assemblage Formation and Archaeological Interpretation in Global Perspective*, University of Colorado Press, Boulder.

Rezek, Z., Lin, S., Iovita, R. and Dibble, H. L. (2011). The Relative Effects of Core Surface Morphology on Flake Shape and Other Attributes. *Journal of Archaeological Science* **38**: 1346–1359.

Richards, M. P., Pettitt, P. B., Trinkaus, E., Smith, F. H., Paunović, M. and Karavanić, I. (2000). Neanderthal Diet at Vindija and Neanderthal Predation: The Evidence from Stable Isotopes. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 7663–6.

Richards, M. P., Taylor, G., Steele, T., McPherron, S. P., Soressi, M., Jaubert, J., Orschiedt, J., Mallye, J. B., Rendu, W. and Hublin, J. J. (2008). Isotopic Dietary Analysis of a Neanderthal and Associated Fauna from the Site of Jonzac (Charente-Maritime), France. *Journal of Human Evolution* **55**: 179–85.

Richards, M. P. and Trinkaus, E. (2009). Out of Africa: Modern Human Origins Special Feature: Isotopic Evidence for the Diets of European Neanderthals and Early

Modern Humans. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 16034–9.

Richter, D., Dibble, H. L., Goldberg, P., McPherron, S. P., Niven, L., Sandgathe, D. M., Talamo, S. and Turq, A. (2013). The Late Middle Palaeolithic in Southwest France: New TL Dates for the Sequence of Pech de l'Azé IV. *Quaternary International* **294**: 160–167.

Van Riper, A. B. (1993). *Men Among the Mammoths: Victorian Science and the Discovery of Human Prehistory*, University of Chicago Press, Chicago.

Rivals, F., Schulz, E. and Kaiser, T. M. (2009). Late and Middle Pleistocene Ungulates Dietary Diversity in Western Europe Indicate Variations of Neanderthal Paleoenvironments through Time and Space. *Quaternary Science Reviews* **28**: 3388–3400.

Rolland, N. and Dibble, H. L. (1990). A New Synthesis of Middle Paleolithic Variability. *American Antiquity* **55**: 480–499.

Rossano, M. J. (2003). Expertise and the Evolution of Consciousness. *Cognition* **89**: 207–236.

Rossano, M. J. (2009). The Archaeology of Consciousness. In S. A. de Beaune F. L. Coolidge and T. Wynn (eds.), *Cognitive Archaeology and Human Evolution*, Cambridge University Press, Cambridge, pp.25–36.

Roth, B. J. and Dibble, H. L. (1998). Production and Transport of Blanks and Tools at the French Middle Paleolithic Site of Combe- Capelle Bas. *American Antiquity* **63**: 47–62.

Rots, V. and Plisson, H. (2014). Projectiles and the Abuse of the Use-Wear Method in a Search for Impact. *Journal of Archaeological Science* **48**: 154–165.

Roussel, M. (2013). Méthodes et Rythmes du Débitage Laminaire au Châtelperronien: Comparaison avec le Protoaurignacien. *Comptes Rendus Palevol* **12**: 233–241.

Sahle, Y., Hutchings, W. K., Braun, D. R., Sealy, J. C., Morgan, L. E., Negash, A. and Atnafu, B. (2013). Earliest Stone-Tipped Projectiles from the Ethiopian Rift Date to >279,000 years ago. *PloS one* **8**: e78092.

Sahnouni, M., Schick, K. D. and Toth, N. (1997). An Experimental Investigation into the Nature of Faceted Limestone "Spheroids" in the Early Palaeolithic. *Journal of Archaeological Science* **24**: 701–713.

Salazar-García, D. C., Power, R. C., Sanchis Serra, A., Villaverde, V., Walker, M. J. and Henry, A. G. (2013). Neanderthal Diets in Central and Southeastern Mediterranean Iberia. *Quaternary International* **318**: 3–18.

Salmon, M. H. (1975). Confirmation and Explanation in Archaeology. *American Antiquity* **40**: 459–464.

Sandgathe, D. M. (2004). Alternative Interpretation of the Levallois Reduction Technique. *Lithic Technology* **29**: 147–159.

Sandgathe, D. M., Dibble, H. L., Goldberg, P. and McPherron, S. P. (2011a). The Roc de Marsal Neandertal Child: A Reassessment of its Status as a Deliberate Burial. *Journal of Human Evolution* **2011**: 1–11.

Sandgathe, D. M., Dibble, H. L., Goldberg, P., McPherron, S. P., Turq, A., Niven, L. and Hodgkins, J. (2011b). On the Role of Fire in Neanderthal Adaptations in Western Europe: Evidence from Pech de l'Aze IV and Roc de Marsal, France. *PaleoAnthropology* **2011**: 216–242.

Sandgathe, D. M., Dibble, H. L., Goldberg, P., McPherron, S. P., Turq, A. and Schwortz, S. (2008). *Roc de Marsal (Campagne-de- Bugue, Dordogne). Rapport d'Opération pour l'Année 2008*, Unpublished report, Bordeaux.

Schick, K. D., Toth, N., Garufi, G., Savage-Rumbaugh, E. S., Rumbaugh, S. M. and Sevcik, R. A. (1999). Continuing Investigations into the Stone Tool-making and Tool-using Capabilities of a Bonobo (Pan paniscus). *Journal of Archaeological Science* **26**: 821–832.

Schiffer, M. B. (1975). Archaeology as Behavioral Science. *American Anthropologist* **77**: 836–848.

Schiffer, M. B. (1998). The Structure of Archaeological Theory. *American Antiquity* **53**: 461–485.

Schlanger, N. (1992). Understanding Levallois: Lithic Technology and Cognitive Archaeology. *Cambridge Archaeological Journal* **6**: 231–254.

Schoville, B. J. (2010). Frequency and Distribution of Edge Damage on Middle Stone Age Lithic Points, Pinnacle Point 13B, South Africa. *Journal of Human Evolution* **59**: 378–391.

Scimelmitz, R., Barkai, R. and Gopher, A. (2011). Systematic Blade Production at Late Lower Paleolithic (400–200 kyr) Qesem Cave, Israel. *Journal of Human Evolution* **61**: 458–479.

Sellet, F. (1993). Chaîne opératoire: The Concept and its Applications. *Lithic Technology* **18**: 106–112.

Semenov, S. A. (1964). *Prehistoric Technology*, Cory, Adams and MacKay, London.

Shackley, M. S. (1998). Gamma Rays, X-Rays and Stone Tools: Some Recent Advances in Archaeological Geochemistry. *Journal of Archaeological Science* **25**: 259–270.

Shea, J. J. (1988). Spear Points from the Middle Paleolithic of the Levant. *Journal of Field Archaeology* **15**: 11–17.

Shea, J. J. (1993). Lithic Use-Wear Evidence for Hunting by Neandertals and Early Modern Humans from the Levantine Mousterian. *Archaeological Papers of the American Anthropological Association* **4**: 189–197.

Shea, J. J. (2006). The Origins of Lithic Projectile Point Technology: Evidence from Africa, the Levant, and Europe. *Journal of Archaeological Science* **33**: 823–846.

Shea, J. J. (2011). Stone Tool Analysis and Human Origins Research: Some Advice from Uncle Screwtape. *Evolutionary Anthropology: Issues, News, and Reviews* **20**: 48–53.

Shea, J. J. and Sisk, M. L. (2010). Complex Projectile Technology and Homo sapeins Dispersal into Western Eurasia. *PaleoAnthropology* **2010**: 100–122.

Sheets, P. D. and Muto, G. R. (1972). Pressure Blades and Total Cutting Edge: An Experiment in Lithic Technology. *Science* **175**: 632–4.

Shennan, S. J. (2000). Population, Culture History, and the Dynamics of Culture Change. *Current Anthropology* **41**: 811–835.

Shennan, S. J. (2008). Evolution in Archaeology. *Annual Review of Anthropology* **37**: 75–91.

Shennan, S. J. (2011). Descent with Modification and the Archaeological Record. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **366**: 1070–1079.

Sheppard, P. J., Irwin, G. J., Lin, S. C. and McCaffrey, C. P. (2011). Characterization of New Zealand Obsidian Using PXRF. *Journal of Archaeological Science* **38**: 45–56.

Shott, M. J. (1986). Technological Organization and Settlement Mobility: An Ethnographic Examination. *Journal of Anthropological Research* **42**: 15–51.

Shott, M. J. (1989). On Tool-Class Use Lives and the Formation of Archaeological Assemblages. *American Antiquity* **54**: 9–30.

Shott, M. J. (1996a). An Exegesis of the Curation Concept. *Journal of Anthropological Archaeology* **52**: 259–280.

Shott, M. J. (1996b). Stage Versus Continuum in the Debris Assemblage from Production of a Fluted Biface. *Lithic Technology* **21**: 6–22.

Shott, M. J. and Sillitoe, P. (2004). Modeling Use-Life Distributions in Archaeology Using New Guinea Wola Ethnographic data. *American Antiquity* **69**: 339–355.

Shott, M. J. and Sillitoe, P. (2005). Use Life and Curation in New Guinea Experimental Used Flakes. *Journal of Archaeological Science* **32**: 653–663.

Sillitoe, P. and Hardy, K. (2003). Living Lithics: Ethnoarchaeology in Highland Papua New Guinea. *Antiquity* **77**: 555–566.

Sisk, M. L. and Shea, J. J. (2009). Experimental Use and Quantitative Performance Analysis of Triangular Flakes (Levallois Points) Used as Arrowheads. *Journal of Archaeological Science* **36**: 2039–2047.

Sisk, M. L. and Shea, J. J. (2011). The African Origin of Complex Projectile Technology: An Analysis Using Tip Cross-Sectional Area and Perimeter. *International Journal of Evolutionary Biology* **2011**: 968012.

Skertchly, S. (1879). *On the Manufacture of Gunflints; The Methods of Excavating for Flint: The Connection between Neolithic Art and the Gunflint Trade.*, District Memoir of the Geological Survey of Great Britain and Ireland, London.

Smith, E. A. (1979). Human Adaptation and Energetic Efficiency. *Human Ecology* **7**: 53–74.

Soffer, O. (2004). Recovering Perishable Technologies through Use Wear on Tools: Preliminary Evidence for Upper Paleolithic Weaving and Net Making. *Current Anthropology* **45**: 407–413.

Sollberger, J. B. (1985). A Technique For Folsom Fluting. *Lithic Technology* **14**: 41–50.

Sollberger, J. B. and Patterson, L. W. (1976). Prismatic Blade Replication. *American Antiquity* **41**: 517–531.

Soressi, M. (2002). *Le Moustérien de Tradition Acheuléenne du Sud-Ouest de la France. Discussion sur la Signification du Faciès à Partir de l'Étude Comparée de Quatre*

Sites: Pech-de-l'Azé I, Le Moustier, La Rochette et la Grotte XVI, unpublished Ph.D. thesis, Université de Bordeaux I

Spaulding, A. C. (1953). Statistical Techniques for the Discovery of Artifact Types. *American Antiquity* **18**: 305–313.

Spaulding, A. C. (1954). Reply to Ford. *American Antiquity* **19**: 391–393.

Speth, J. D. (1972). The Mechanical Basis of Percussion Flaking. *American Antiquity* **37**: 34–60.

Speth, J. D. (1974). Experimental Investigations of Hard-Hammer Percussion Flaking. *Tebiwa* **17**: 7–36.

Speth, J. D. (1981). The Role of Platform Angle and Core Size in Hard-Hammer Percussion Flaking. *Lithic Technology* **10**: 16–21.

Stern, N. (1994). The Implications of Time-Averaging for Reconstructing the Land-Use Patterns of Early Tool-Using Hominids. *Journal of Human Evolution* **27**: 89–105.

Stiles, D. (1977). Ethnoarchaeology : A Discussion of Methods and Applications. *Man, New Series* **12**: 87–103.

Stout, D. (2011). Stone Toolmaking and the Evolution of Human Culture and Cognition. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences* **366**: 1050–9.

Stout, D., Apel, J., Commander, J. and Roberts, M. (2014). Late Acheulean Technology and Cognition at Boxgrove, UK. *Journal of Archaeological Science* **41**: 578–590.

Stout, D. and Chaminade, T. (2009). Making Tools and Making Sense: Complex, Intentional Behaviour in Human Evolution. *Cambridge Archaeological Journal* **19**: 85–96.

Stout, D. and Chaminade, T. (2012). Stone Tools, Language and the Brain in Human Evolution. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences* **367**: 75–87.

Stout, D., Passingham, R., Frith, C., Apel, J. and Chaminade, T. (2011). Technology, Expertise and Social Cognition in Human Evolution. *The European Journal of Neuroscience* **33**: 1328–38.

Stout, D., Toth, N., Schick, K. and Chaminade, T. (2008). Neural Correlates of Early Stone Age Toolmaking: Technology, Language and Cognition in Human Evolution.

Philosophical transactions of the Royal Society of London. Series B, Biological Sciences **363**: 1939–49.

Stout, D., Toth, N., Schick, K., Stout, J. and Hutchins, G. (2000). Stone Tool-Making and Brain Activation: Position Emission Tomography (PET) Studies. *Journal of Archaeological Science* **27**: 1215–1223.

Sullivan, A. P. (1978). Inference and Evidence in Archaeology : A Discussion of the Conceptual Problems. *Advances in Archaeological Method and Theory* **1**: 183–222.

Tactikos, J. C. (2003). A Re-Evaluation of Palaeolithic Stone Tool Cutting Edge Production Rates and Their Implications. In N. Moloney and M. J. Shott (eds.), *Lithic Analysis in the Millenium*, Left Coast Press, London, pp.151–162.

Tennie, C., Call, J. and Tomasello, M. (2009). Ratcheting Up the Ratchet: On the Evolution of Cumulative Culture. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences* **364**: 2405–15.

Thieme, H. (1997). Lower Palaeolithic Hunting Spears from Germany. *Nature* **385**: 807–10.

Tite, M. S. (1991). Archaeological Science - Past Achievements and Future Prospects. *Archaeometry* **33**: 139–151.

Tixier, J., Inizan, M.-L. and Roche, H. (1980). *Préhistoire de la Pierre Taillé. Vol. 1: Terminologie et technologie*, Cercle de Recherches et d'Etudes Préhistorique, Antibes.

Torrence, R. (1983). Time Budgeting and Hunter-Gatherer Technology. In G. Bailey (ed.), *Hunter-Gatherer Economy in Prehistory: A European Perspective*, Cambridge University Press, Cambridge, pp.11–22.

Torrence, R. ed. (1989). *Time, Energy, and Stone Tools*, Cambridge University Press, Cambridge.

Torrence, R. (2001). Hunter-Gatherer Technology: Macro- and Microscale Approaches. In C. Panter-Brick R. H. Layton and P. Rowley-Conwy (eds.), *Hunter-Gatherers: An Interdisciplinary Perspective*, Cambridge University Press, Cambridge, pp.73–99.

Tostevin, G. B. (2013). *Seeing Lithics: A Middle-Range Theory for Testing for Cultural Transmission in the Pleistocene*, Oxbow Books, Oakville.

Toth, N., Schick, K. D., Savage-Rumbaugh, E. S., Sevcik, R. A. and Rumbaugh, S. M. (1993). Pan the Tool-Maker: Investigations into the Stone Tool-Making and Tool-Using Capabilities of a Bonobo (Pan paniscus). *Journal of Archaeological Science* **20**: 81–91.

Trigger, B. G. (1995). Expanding Middle-Range Theory. *Antiquity* **69**: 449–458.

Trigger, B. G. (2006). *A History of Archaeology Thought*, Cambridge University Press, Cambridge.

Tschauner, H. (1996). Middle-Range Theory, Behavioral Archaeology, and Postempiricist Philosophy of Science in Archaeology. *Journal of Archaeological Method and Theory* **3**: 1–30.

Turq, A. (1989). Approche Technologique et Économique du Facies Moustérien de Type Quina: Etude Préliminaire. *Bulletin de la Société Préhistorique Française* **86**: 244–256.

Turq, A. (1992). Raw Material and Technological Studies of the Quina Mousterian in Perigord. In H. L. Dibble and P. Mellars (eds.), *The Middle Paleolithic: Adaptation, Behavior, and Variability*, University Museum Press, Philadelphia, pp.75–85.

Turq, A. (1995). Raw Material Sources in the Region of Combe-Capelle Bas. In H. L. Dibble and M. Lenoir (eds.), *The Middle Paleolithic Site of Combe-Capelle Bas (France*, University Museum Press, Philadelphia, pp.321–328.

Turq, A. (2000). Paléolithique Inférieur et Moyen entre Dordogne et Lot. *Paléo* **2**: 1–456.

Turq, A., Dibble, H. L., Faivre, J., Goldberg, P., McPherron, S. P. and Sandgathe, D. M. (2008). Le Moustérien du Périgord Noir: quoi de neuf? In J. Jaubert J.-G. Bordes and I. Ortega (eds.), *Les Sociétés Paléolithiques D'un Grand Sud-Ouest: Nouveaux Gisements, Nouvelles Méthodes, Nouveaux Résultats,*, Bulletin de la Societe Prehistorique Francaise, Paris, pp.83–94.

Turq, A., Dibble, H. L., Goldberg, P., McPherron, S. P., Sandgathe, D. M., Jones, H., Maddison, K., Maureille, B., Mentzer, S., Rink, J. and Steenhuyse, A. (2011). Les Fouilles Récentes du Pech de l'Azé IV (Dordogne). *Gallia Préhistoire* **53**: 1–58.

Turq, A., Roebroeks, W., Bourguignon, L. and Faivre, J.-P. (2013). The Fragmented Character of Middle Palaeolithic Stone Tool Technology. *Journal of Human Evolution* **65**: 641–55.

Valladas, H., Mercier, N., Joron, J. L., McPherron, S. P., Dibble, H. L. and Lenoir, M. (2003). TL dates for the Middle Paleolithic site of Combe-Capelle Bas, France. *Journal of Archaeological Science* **30**: 1443–1450.

Vandenberghe, J. (1995). Timescales, Climate and River Development. *Quaternary Science Reviews* **14**: 631–638.

Vandenberghe, J. (2001). A Typology of Pleistocene Cold-Based Rivers. *Quaternary International* **79**: 111–121.

Vandenberghe, J. (2002). The Relation between Climate and River Processes, Landforms and Deposits during the Quaternary. *Quaternary International* **91**: 17–23.

Vandenberghe, J. (2003). Climate Forcing of Fluvial System Development: An Evolution of Ideas. *Quaternary Science Reviews* **22**: 2053–2060.

Vandenberghe, J. (2008). The Fluvial Cycle at Cold–Warm–Cold Transitions in Lowland Regions: A Refinement of Theory. *Geomorphology* **98**: 275–284.

Vaquero, M. (2008). The History of Stones: Behavioural Inferences and Temporal Resolution of an Archaeological Assemblage from the Middle Palaeolithic. *Journal of Archaeological Science* **35**: 3178–3185.

Vaquero, M., Chacón, M. G., García-Antón, M. D., Gómez de Soler, B., Martínez, K. and Cuartero, F. (2012). Time and Space in the Formation of Lithic Assemblages: The Example of Abric Romaní Level J. *Quaternary International* **247**: 162–181.

Villa, P. and Lenoir, M. (2006). Hunting Weapons of the Middle Stone Age and the Middle Palaeolithic: Spear Points from Sibudu, Rose Cottage and Bouheben. *Southern African Humanities* **18**: 89–122.

Wadley, L. (2005). Putting Ochre to the Test: Replication Studies of Adhesives that May Have Been Used for Hafting Tools in the Middle Stone Age. *Journal of Human Evolution* **49**: 587–601.

Wadley, L. (2010). Compound-Adhesive Manufacture as a Behavioral Proxy for Complex Cognition in the Middle Stone Age. *Current Anthropology* **51**: S111–S119.

Wadley, L., Williamson, B. and Lombard, M. (2004). Ochre in Hafting in Middle Stone Age Southern Africa: A Practical Role. *World Archaeology* **26**: 19–34.

Walker, P. L. (1978). Butchering and Stone Tool Function. *American Antiquity* **43**: 710–715.

Wallace, I. J. and Shea, J. J. (2006). Mobility Patterns and Core Technologies in the Middle Paleolithic of the Levant. *Journal of Archaeological Science* **33**: 1293–1309.

Wandsnider, L. (2004). Solving Meno's Puzzle, Defeating Merlin's Subterfuge: Bodies of Reference Knowledge and Archaeological Inference. In A. L. Johnson (ed.), *Processual Archaeology: Exploring Analytical Strategies, Frames of Reference, and Culture Process*, Greenwood Press, pp.275–305.

Watson, P. J., LeBlanc, S. A. and Redman, C. L. (1971). *Explanation in Archaeology: An Explicitly Scientific Approach*, Columbia University Press, New York.

White, J. P. (1967). Ethno-Archaeology in New Guinea: Two Examples. *Mankind* **6**: 409–414.

White, J. P. and Thomas, D. H. (1972). What Mean these Stones? Ethno-Taxonomic Models and Archaeological Interpretations in the New Guinea Highlands. In D. Clarke (ed.), *Models in Archaeology*, Methuen Publishing Ltd, London, pp.275–308.

White, P. and White, C. (1964). Surface Sites in South Africa. *The South African Archaeological Bulletin* **19**: 64–66.

Whittaker, J. C. (1994). *Flintknapping: Making & Understanding Stone Tools*, University of Texas Press, Austin.

Wilkins, J., Schoville, B. J. and Brown, K. S. (2014). An Experimental Investigation of the Functional Hypothesis and Evolutionary Advantage of Stone-Tipped Spears. *PloS one* **9**: e104514.

Wilkins, J., Schoville, B. J., Brown, K. S. and Chazan, M. (2012). Evidence for Early Hafted Hunting Technology. *Science* **338**: 942–6.

Williams, E. M. (2011). *Biomechanical Strategies during Oldowan and Acheulean Stone Tool Production*, unpublished Ph.D. dissertation, The George Washington University.

Williams, E. M., Gordon, A. D. and Richmond, B. G. (2012). Hand Pressure Distribution during Oldowan Stone Tool Production. *Journal of Human Evolution* **62**: 520–32.

Winterhalder, B. (2001). The Behavioral Ecology of Hunter-Gatherers. In C. Panter-Brick R. H. Layton and P. Rowley-Conwy (eds.), *Hunter-Gatherers: An Interdisciplinary Perspective*, Cambridge University Press, Cambridge, pp.13–38.

Wylie, A. (1982). Epistemological Issues Raised by a Structuralist Archaeology. In I. Hodder (ed.), *Symbolic and Structural Archaeology*, Cambridge University Press, Cambridge, pp.39–46.

Wylie, A. (1985). The Reaction Against Analogy. *Advances in Archaeological Method and Theory* **8**: 63–111.

Wylie, A. (1993). A Proliferation of New Archaeologies: "Beyond Objectivism and Relativism". In N. Yoffee and A. Sherratt (eds.), *Archaeological Theory: Who Sets the Agenda*, Cambridge University Press, Cambridge, pp.20–26.

Wylie, A. (2000). Questions of Evidence, Legitimacy, and the (Dis)Unity of Science. *American Antiquity* **65**: 227–237.

Wynn, T. (1981). The Intelligence of Oldowan Hominids. *Journal of Human Evolution* **10**: 529–541.

Wynn, T. (1995). Handaxe Enigmas. *World Archaeology* **27**: 10–24.

Wynn, T. (2002). Archaeology and Cognitive Evolution. *Behavior and Brain Science* **25**: 389–438.

Wynn, T. (2008). Tools, Grammar and the Archaeology of Cognition. *Cambridge Archaeological Journal* **1**: 191.

Wynn, T. and Coolidge, F. L. (2004). The Expert Neandertal Mind. *Journal of Human Evolution* **46**: 467–87.

Yaroshevich, A., Nadel, D. and Tsatskin, A. (2013). Composite Projectiles and Hafting Technologies at Ohalo II (23 ka, Israel): Analyses of Impact Fractures, Morphometric Characteristics and Adhesive Remains on Microlithic Tools. *Journal of Archaeological Science* **40**: 4009–4023.

Zwyns, N. (2012). *Laminar technology and the onset of the Upper Paleolithic in the Altai, Siberia*, Leiden University Press, Leiden.