New Experimental Insights into Early Hominin Cultures and Oldowan Technology

Dissertation
der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors des Naturwissenschaften
(Dr. rer. nat.)

Vorgelegt von
William Daniel Snyder
aus Burnsville, MN, USA

Tübingen
2022

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation: 27.01.2023

Dekan: Prof. Dr. Thilo Stehle

1. Berichterstatter/-in: Dr. Claudio Tennie

2. Berichterstatter/-in: Prof. Dr. Nicholas Conard

Acknowledgments

Being able to work as a researcher in any scientific field is an incredible and special privilege. Yet, being a researcher nowadays can mean an incredible pressure to finish our degrees, publish article, and find jobs. This pressure can have disastrous effects on the mental wellbeing of the few and the many. Stress and other mental health factors, combined with limited opportunities, lead to the exodus of too many otherwise talented thinkers and empiricists from academia. This dissertation is dedicated to all my friends and colleagues who have struggled during their time in academia and especially to those who have been driven from the career of their dreams for the wrong reasons.

Above all, I want to thank my parents, Michael and Claudia, who have always supported me and nurtured my passions and inquisitive mind. To my sister Leslie, I call attention to the fact that you will soon no longer be the only Dr. in the family. To the whole family, I love you. Opa and Oma, you are dearly missed.

I am thankful for all of my wonderful friends and peers who have shared this time with me in Tübingen over the last seven years. I could have never made it this far without the lot of you.

Without the job offer from Dr. Claudio Tennie, I would never have had the chance to complete this important work. Claudio, thank you so much for giving me the opportunity and entrusting me with such a valuable project. I can only hope that expectations were met.

Likewise, thank you to my other supervisor Prof. Nicholas Conard, to the other members of my defense committee, Prof. Katerina Harvati and Prof. Dietrich Stout, and to the remaining member of my EVEREST advisory committee, Prof. Nico Michiels.

Special thanks to all of my (now former) lab mates and collaborators in and affiliated with STONECULT: Elisa Bandini, David Boysen, Nolan Ferar, Jordy Orellana Figueroa, Li Li, Shannon McPherron, Alba Motes-Rodrigo, Damien Neadle, Jonathan S. Reeves, Eva Reindl, and Vera Thomas.

For their logistical support regarding the testing space, I give my thanks to Angel Blanco Lapaz, Madita Kairies, and Britt Starkovich.

For the work on raw material standardization, I want to express my appreciation for: Silvia Amicone, Patrick Cuthbertson, Sabine Flaiz, Peter Kühn, Sam Lin, Rita Mögenburg, Klaus Nickel, Thomas Nieß, Kathie Ranhorn, Karin Rein, and Cory Stade.

For their service as interreliability coders, I would like to highlight the contributions of Armando Falcucci and James Keppeler.

This work was funded by the European Research Council (ERC) Horizon 2020 research and innovation program (grant agreement n°714658; STONECULT project, C.T.).

Statement on Collaborative Work

This dissertation is the product of collaborative work conducted in the scope of the ERC Starting Grant project STONECULT, supervised by Dr. Claudio Tennie.

Where applicable, the contributions of the candidate and of the co-authors are indicated at the start of a chapter. These contributions are separated into the following basic categories: Conceptualization, Investigation, Visualization, Formal analysis, Writing (original draft), and Writing (review and editing). Authors are listed per category according to their relative contribution to the specified task.

In some cases, only excerpts of a manuscript or publication are included in the relevant chapter. Unless otherwise indicated in the figure caption, figures were prepared by the candidate.

Additionally, this dissertation contains work that is intended to be published but which has not yet been edited by co-authors (thus, representing what is for now the sole work of the candidate).

Candidate Bibliography

The following bibliography represents the sum of work (original research, analyses, and reviews) performed by the candidate during his doctoral studies. Publications or manuscripts that appear either in part or in whole (or are extensively summarized) in the present dissertation are marked with *.

Publications

* **Snyder, W.D.** & Tennie, C. (in press). What kind of culture did early hominin toolmakers have?. *Mitteilungen der Berliner Gesellschaft für Anthropologie, Ethnologie und Urgeschichte*.

* Acerbi, A., **Snyder, W.D.**, & Tennie, C. (2022). The method of exclusion (still) cannot identify mechanisms of cultural inheritance. *Scientific Reports*, *12*, 21680. doi: 10.1038/s41598-022-25646-9

* **Snyder, W.D.**, Reeves, J.S., & Tennie, C. (2022). Early knapping techniques do not necessitate cultural transmission. *Science Advances, 8*(27), eabo2894. doi:10.1126/sciadv.abo2894

Snyder, W.D. (2022). Have video games evolved enough to teach human origins? A review of *Ancestors: The Humankind Odyssey*. *Advances in Archaeological Practice, 10*(1), 122-127. doi:10.1017/aap.2021.40

Bandini, E., Reeves, J.S., **Snyder, W.D.**, & Tennie, C. (2021). Clarifying misconceptions of the Zone of Latent Solutions hypothesis: A response to Haidle & Schlaudt, 2020". *Biological Theory, 16*(5), 76-82. doi:10.1007/s13752-021-00374-x

In Preparation

* **Snyder, W.D.** & Tennie, C. (in prep). Novice knapping experiments: possibilities, pitfalls, and potential in cognitive archaeology.

* **Snyder, W.D.**, Boysen, D., Orellana-Figueroa, J.D., Tennie, C., & Reeves, J.S. (in prep). An overview of standardizable raw materials for controlled human knapper experiments.

* **Snyder, W.D.**, Reeves, J.S., & Tennie, C. (in prep). Untitled naïve human toolmaking and tool use paper.

Abstract

Cumulative culture of know-how is considered a key feature of modern human life, being fundamental to the development and diversification of technology, as well as norms, practices, and beliefs. This phenomenon and the underlying copying social learning mechanisms are likely unique – as far as technological know-how – to living modern humans. Other species also have culture, but their *minimal culture* is guided by non-copying social learning mechanisms that transmit information types like know-what and know-where. Based on the absence of copying social learning mechanisms and cumulative culture of know-how in other species – including non-human great apes – it can be inferred that these mechanisms and processes first emerged during the evolution of the hominin lineage.

Due to taphonomic biases (thus limiting from what cognitive archaeologists can make inferences), most of what is left over from millions of years of hominin behaviors and cognition is a 2.6-million-year record of stone artifacts and other tool-use related end-products. Previous authors – often using data from novice knapping experiments as justification – have concluded that even the earliest products of hominin toolmaking in the Oldowan are evidence that hominins were behaving and *thinking* in ways that are culturally similar to the cumulative culture of know-how in living modern humans. Recent alternative hypotheses have gone against this line of inference, instead proposing – in the same way as for non-human great apes and other primates – that Oldowan technology was a consequence of serial re-innovations and cultural transmission of non-know-how information types.

Cognitive and experimental archaeology as an entire discipline faces a vast swathe of epistemological concerns and considerations. Although novice knapping experiments serve a fundamental purpose in the construction of middle range theories about the human and hominin past, they are often entrenched in pathways of circular reasoning and uncontextualized analogies. Instead of focusing on determining what learning conditions are most efficient or effective for the acquisition of knapping know-how by human participants, cognitive archaeology should return to its root: the determination of minimal necessary competences for the existence of artifact types and knapping techniques.

As such, we adopted a baseline methodological approach (i.e., an island test) for determining whether copying and cultural transmission of know-how are absolutely necessary for the

expression of early knapping techniques by modern human participants. Rather than teach or somehow train the participants to knap, knapping know-how was elicited using the provision of appropriate raw materials and the motivation to use tools (i.e., a baited puzzle box). With our island test approach, we were able to identify manifold occurrences of early knapping technique re-innovations, demonstrating that the know-how for Oldowan toolmaking does not exceed the reach of individual re-innovation capacities and therefore – by proof-of-principle – Oldowan stone artifacts and their requisite manufacturing techniques cannot be considered unequivocal evidence for early cases of cultural transmission of know-how in the hominin lineage. The major theoretical outcome of this study – and outlined in this dissertation – is a minimal culture model for premodern hominins and technological change in the early stone tool record.

In this dissertation, I introduce concepts that are fundamental to discourse on non-human cognition as well as the evolution of cognition and technology in the human lineage (Chapter I), review the breadth of novice knapping experiments that have been conducted so far and the epistemological issues that are inherent to this avenue of research (Chapter II), outline the different raw materials and standardization techniques that we tested for use in knapping experiments (Chapter III), describe the methodology that we applied in further detail (Chapter IV), reveal the most important results that came out of the experiments (Chapter V), contextualize the results based on previous and ongoing research related to culture in nonhumans and hominins (Chapter VI), and – finally – discuss the theoretical implications of these and other available data for a minimal culture model of the Oldowan along with other aspects of hominin cognitive and technological evolution.

Abstrakt

Die kumulative Kulter des Know-hows gilt als ein wesentliches Merkmal des modernen Menschen, da sie für die Entwicklung und Diversifizierung von Technologie sowie von Normen, Praktiken, und Überzeugungen von grundlegender Bedeutung ist. Dieses Phänomen und die zugrundeliegenden Mechanismen des kopierenden sozialen Lernens sind - was das technologische Know-how betrifft - wahrscheinlich einzigartig für lebende moderne Menschen. Andere Tierarten haben zwar auch eine Kultur, aber ihre Minimal-Kultur wird durch nicht-kopierende soziale Lernmechanismen gesteuert, die Informationen wie Know-what und Know-where weitergeben. Aus der Abwesenheit kopierender sozialer Lernmechanismen und einer kumulativen Kultur des Know-hows bei anderen Tierarten - einschließlich nichtmenschlicher Menschenaffen - lässt sich ableiten, dass diese Mechanismen und Prozesse erst während der Evolution der Linie von der Hominen entstanden sind.

Aufgrund von taphonomischen Verzerrungen (die die Möglichkeiten der kognitiven Archäologen einschränken) ist das meiste, was von Millionen von Jahren homininen Verhaltens und kognitiver Fähigkeiten übrig geblieben ist, einen 2,6-millionenjährigen Rekord von Steinartefakten und anderen mit dem Werkzeuggebrauch verbundenen Endprodukten. Frühere Autoren haben - oft unter Verwendung von Daten aus Experimenten mit Anfängern beim Steinschlagen - den Schluss gezogen, dass selbst die frühesten Produkte der homininen Werkzeugherstellung im Oldowan ein Beweis dafür sind, dass Homininen in einer Art und Weise handelten und dachten, die der kumulativen Kultur des Know-hows bei lebenden modernen Menschen kulturell ähnlich ist. Neuere alternative Hypothesen widersprechen dieser Schlussfolgerung und schlagen stattdessen vor, dass die Technologie des Oldowan eine Folge serieller Re-Innovationen und kultureller Weitergabe von Informationsarten war, die nicht zum Know-How gehören, wie dies bei Menschenaffen und anderen Primaten der Fall ist.

Die kognitive und experimentelle Archäologie ist als gesamte Disziplin mit einer Vielzahl von erkenntnistheoretischen Bedenken und Überlegungen konfrontiert. Obwohl Steinschlagsexperimente für Anfänger einen grundlegenden Zweck bei der Konstruktion von Theorien über die menschliche und hominine Vergangenheit erfüllen, sind sie oft in Zirkelschlüssen und unkontextualisierten Analogien verhaftet. Anstatt sich auf die Frage zu

konzentrieren, welche Lernbedingungen am effizientesten oder effektivsten für den Erwerb von Steinschlagskenntnissen durch menschliche Teilnehmer sind, sollte die kognitive Archäologie zu ihren Wurzeln zurückkehren: die Bestimmung der minimal notwendigen Kompetenzen für die Existenz von Artefakttypen und Steinschlagstechniken.

Daher haben wir einen grundlegenden methodischen Ansatz (d. h. einen Inseltest) gewählt, um festzustellen, ob das Kopieren und die kulturelle Weitergabe von Know-how für die Ausprägung frühe Steinschlagstechniken durch moderne menschliche Teilnehmer absolut notwendig sind. Anstatt den Teilnehmern das Steinschlagen beizubringen oder sie in irgendeiner Weise zu trainieren, wurde das Know-how durch die Bereitstellung geeigneter Rohstoffe und die Motivation zur Verwendung von Werkzeugen (z. B. eine mit einem Köder versehene Rätselkiste) erlangt. Mit unserem Inseltest-Ansatz waren wir in der Lage, vielfältige Vorkommnisse Re-Innovationen von frühe Steinschlagstechniken zu identifizieren, was zeigt, dass das Know-how für die Herstellung von altweltlichen Werkzeugen die Reichweite individueller Re-Innovationskapazitäten nicht übersteigt und daher - als proof-of-principle - altweltliche Steinartefakte und die dazugehörigen Herstellungstechniken nicht als eindeutiger Beweis für frühe Fälle kultureller Weitergabe von Know-how in der Hominінen-Linie gelten können. Das wichtigste theoretische Ergebnis dieser Studie - und in dieser Dissertation dargelegt - ist ein Minimal-Kultur-Modell für vormoderne Homininen und technologischen Wandel in den frühen Steinwerkzeugfunden.

In dieser Dissertation führe ich Konzepte ein, die für den Diskurs über nicht-menschliche Kognition sowie die Entwicklung von Kognition und Technologie in der menschlichen Abstammungslinie von grundlegender Bedeutung sind (Kapitel I), gebe einen Überblick über die Bandbreite der bisher durchgeführten Steinschlagsexperimente mit Anfängern und die erkenntnistheoretischen Fragen, die mit diesem Forschungszweig verbunden sind (Kapitel II), beschreibe die verschiedenen Rohmaterialien und Standardisierungstechniken, die wir für die Verwendung in Steinschlagsexperimenten getestet haben (Kapitel III), beschreibe die von uns angewandte Methodik im Detail (Kapitel IV), zeige die wichtigsten Ergebnisse der Experimente auf (Kapitel V), stelle die Ergebnisse in den Kontext früherer und laufender Forschungen zur Kultur von Nichtmenschen und Homininen (Kapitel VI), und diskutiere schließlich die theoretischen Implikationen dieser und anderer verfügbarer Daten für ein Minimal-Kultur-Modell des Oldowan und andere Aspekte der kognitiven und technologischen Evolution von Homininen.

Table of Contents

List of Figures

List of Tables

I. Introduction

Explanation of candidate contribution to collaborative work

This chapter is based on collaborative work. The text is partially adapted from the main text of the following publication:

Snyder, W.D., Reeves, J.S., & Tennie, C. (2022). Early knapping techniques do not necessitate cultural transmission. *Science Advances*, *8*(27), eabo2894. doi:10.1126/sciadv.abo2894

Conceptualization: C.T., W.D.S., J.S.R. Investigation: W.D.S. Visualization: W.D.S. Formal analysis: W.D.S., C.T. Writing (original draft): W.D.S. Writing (review and editing): W.D.S., C.T.

Some of the remaining content is from as-yet unpublished, in-preparation work:

Snyder, W.D. & Tennie, C. (in press). What kind of culture did early hominin toolmakers have?. *Mitteilungen der Berliner Gesellschaft für Anthropologie, Ethnologie und Urgeschichte*.

Snyder, W.D., Reeves, J.S., & Tennie, C. (in prep). Untitled naïve human toolmaking and tool use paper.

From the word *culture* arises a vast array of connotations and associations related to existential and social notions of who we are as individuals and as groups, what makes us human, and what separates us from all the other diverse and – each in their own way – unique life forms that share this planet. Just as it has exceptional philosophical value, culture is a central theme for a great number of discussions and debates in academia and is the subject of manifold studies across scientific disciplines (e.g., Boyd, 2018; Boyd & Richerson, 1996; Dean et al., 2013; Henrich, 2016; Heyes, 2020, 2021; Tennie et al., 2020; Tomasello, 1999; Watson, 1995), including – but by no means limited to – anthropology, archaeology, linguistics, psychology, cognitive science, neuroscience, digital humanities, and research into non-human animal behavior and proposed traditions (e.g., Dean et al., 2012, 2013; Heyes, 2020, 2021; Kurzban & Barrett, 2012; Tennie et al., 2020; Watson, 1995).

Culture conundrums

The matter of defining culture is a rather complicated one, with many disciplines and even singular research groups having their own idiosyncratic – and too often, nonpragmatic – definitions for the term (see Foley & Lahr, 2003; Heyes, 2020; McGrew, 1992; Watson, 1995). In anthropology, culture can refer to the sum of traditions, practices, rites, beliefs, and material associated with a specific human population or smaller group (e.g., Tylor, 1871; Watson, 1995). In archaeology (especially culture-historical archaeology), traces of human behavior and artifactual patterns (e.g., specific artifact typologies or assemblage compositions) are identified as potentially corresponding to groups with other shared characters (e.g., Childe, 1929; discussed in Foley & Lahr, 2003; Watson, 1995).

These more anthropocentric definitions, however, fail to successfully accommodate socially learned behaviors and localized or regionalized behavioral traditions in non-human animals (e.g., Dean et al., 2013; Foley & Lahr, 2003; Heyes, 2020, 2021; McGrew, 1992). Similar difficulties can be found with assigning ‘cultures’ to long-lasting Early Stone Age/Lower Paleolithic artifact types that cannot be so easily assigned to smaller categorizations of local traditions, let alone to any assumed sets of behavioral practices and/or mental beliefs (cf. Davidson & McGrew, 2005; Davidson & Noble, 1993; Delagnes & Roche, 2005; Foley & Lahr, 2003; Tennie et al., 2017; Watson, 1995). In order to view cultural capacities in a large-scale evolutionary framework, it is necessary to adopt more purposeful, mechanism-based definitions of culture and the different kinds of culture (here, divided mainly into the two categories of minimal culture and cumulative culture of know-how; see below). In this sense

– and for the scientific purposes of this dissertation – culture is defined simply as any behavior or information that is under *some kind* of social influence (cf. Aplin, 2016; Boyd & Richerson, 1988; Dean et al., 2013; Foley & Lahr, 2003; Heyes, 2020, 2021; Neadle et al., 2017).

Social and individual learning

Social learning includes a broad range of cognitive mechanisms related to "learning about other agents or the inanimate world that is influenced by observation of, or interaction with, another individual or its products" (Heyes, 2012, p. 193). Usually, the other individual from which an agent learns is a conspecific (Heyes, 1994). For learning whereby there is no influence from either "observation of, or interaction with, another individual or its products" (Heyes, 2012, p. 193), terminology such as individual learning, asocial learning, or trial-and-error learning are typically applied (hereafter, the term individual learning will most often be applied; Heyes, 1994, 2012).

Social learning mechanisms have often been separated into two main categories: high-fidelity social learning and low-fidelity social learning (Lewis & Laland, 2012; Tennie et al., 2009). The terms high-fidelity and low-fidelity social learning refer to the degree of faithfulness with which the form of a particular target behavior is reproduced by an 'observer' (cf. Lewis & Laland, 2012). This would potentially infer that some degree of form copying or cultural transmission of know-how is still occurring via low-fidelity social learning mechanisms, which is not strictly true, but may yet be (see Tennie et al., 2009, 2020). As such – instead of high-fidelity and low-fidelity social learning – the categories of copying social learning and non-copying social learning have been recommended instead (Tennie et al., 2020).

Copying social learning mechanisms are those capable of transmitting *know-how*[1] information (i.e., information within the mind related to the form of either behavior or behavioral outcomes, including artifacts) between and across individuals (thus, *cultural transmission of know-how*), and they include, e.g., emulation, imitation, (certain types of) teaching, and language (e.g., Boyd & Richerson, 1985, 1995; Caldwell & Millen, 2009; Dean et al., 2012; Lewis & Laland, 2012; Reindl et al., 2017; Tennie et al., 2009, 2020; Tomasello, 1999; Tomasello et al., 1993). Though some authors refer to *emulation* (also known as results

[1] Other authors use the term 'know-how' to refer to an overlapping, yet distinct, concept, which might otherwise be referred to as procedural knowledge (cf. Bamforth & Finlay, 2008; Pargeter et al., 2020).

copying: Tennie et al., 2010, 2012; or end-state emulation: e.g., Whiten et al., 2009a) as a 'low-fidelity' social learning mechanism (cf. T. Morgan et al., 2015; Sterelny, 2021), the observation of behavioral outcomes can also be enough to transmit the know-how that produces those outcomes (Caldwell & Millen, 2009; Reindl et al., 2017; contra Boyd & Richerson, 1985; Tomasello et al., 1993; see also Sterelny, 2021) and certainly allows for the copying of artifact forms (see Tennie et al., 2020). *Imitation* typically involves copying of the specific (i.e., topographically matching or similar; Heyes, 1994; Zuberbühler et al., 1996; Whiten et al., 2009a) form of a behavioral trait (hence, action or form copying; Bandini & Tennie, 2019; Bandini et al., 2020; Neadle et al., 2017; Tennie et al., 2020). *Teaching* can refer to any case in which an agent (i.e., a 'teacher') modifies its behavior in some way to facilitate learning in at least one other individual; teaching of know-how would thus involve behaviors like action demonstrations that are then followed by replication of the demonstrated actions. While the expression of language (i.e., language separated out into many individual units; e.g., Motes-Rodrigo & Tennie, 2021) is under the influence of cumulative cultural evolution of know-how, language itself (as an 'advanced' communication capacity) is also a mechanism by which know-how can be transmitted (further, it is often specifically mentioned as a mechanism to directly transfer technical knowledge; Dean et al., 2012, 2013; T. Morgan et al., 2015; Lombao et al., 2017; Cataldo et al., 2018).

Non-copying social learning mechanisms are those capable of transmitting information other than know-how (Tennie et al., 2020), such as know-what (e.g., objects or target organisms), know-where (e.g., positions or geographic locations), know-when (e.g., moments or seasons), and know-who (e.g., parents or unfamiliar conspecifics) (see Bandini et al., 2020; Tennie et al., 2020. Some of the mechanisms responsible for *cultural transmission of know-*x (where know-*x* is an information type *other than* know-how) include, e.g., stimulus enhancement, local enhancement, teaching of know-*x*, exposure, and social contagions (only the former three are discussed in more detail here; see also Tennie et al., 2020). *Stimulus enhancement* is (at least one mechanism of) cultural transmission of know-what. In stimulus enhancement, the interaction of another individual with a specific stimulus increases the likelihood that an individual will themselves interact with that same stimulus (Arbilly & Laland, 2014; Heyes, 1994; Zuberbühler et al., 1996). *Local enhancement* is (at least one mechanism of) cultural transmission of know-where. In local enhancement, the presence of another individual at a specific location increases the likelihood that an individual will visit that place and/or interact

Table 1 Examples of social and individual learning mechanisms according to the information types (i.e., know-what, know-where, know-when, know-who, and know-how) that that they transmit or of which they are instrumental to development.

	‘Transmitted’ Information Types				
	Know-what	Know-where	Know-when	Know-who	Know-how
Social	e.g., stimulus enhancement	e.g., local enhancement	e.g., emergence times*	e.g., mate-choice ‘copying’**	e.g., emulation, imitation, verbal and gestural communication, teaching of know-how
Social	e.g., teaching of know-*x* (i.e., simple scaffolding)				
Individual	e.g., associative learning, trial-and-error learning, conditioning				Individual re-innovation of know-how (i.e., built upon other individual learning processes)

* Meerkats have been shown to have group-specific timing for when they emerge from their sleeping burrows (Thornton et al., 2010)
** One example of know-who transmission would be the ‘copying’ of mate choice, as has been found and used to demonstrate cultural capacities in fruit flies (Danchin et al., 2018).

with objects located there (Arbilly & Laland, 2014; Heyes, 1994; Hoppitt & Laland, 2008). As inferred previously, teaching does not necessarily involve the transmission of know-how, but instead can involve, e.g., the facilitation of learning by provisioning of materials relevant to a behavior (simple scaffolding or teaching of know-what; cf. Thornton & Raihani, 2010).

Both stimulus enhancement and local enhancement are contingent upon associative learning processes (Heyes, 1994), the very same associative learning processes that also guide individual learning mechanisms (Heyes & Pearce, 2015). Association of a learning model (again, whether a conspecific, group member, or someone or something else) with a stimulus or location need “not involve conscious or unconscious reasoning” (Heyes & Pearce, 2015, p. 6), but rather can be inherited via genetics alone. Even further, all social learning has been suggested to share the same mechanistic learning foundations as individual learning (Heyes, 2012). By this token, the role of social learning – in at least some contexts (e.g., as suggested for stone toolmaking in living modern humans; Snyder et al., 2022) – is to expedite know-how development, which might otherwise be exclusively individually-developed.

An alternative pathway for know-how development that does not require copying social learning is individual re-innovation (i.e., copying-independent; Tennie et al., 2020). Unlike innovation, which involves the introduction of new behaviors or modifications to existing behavioral traits and paradigms (Reader & Laland, 2003), a reinnovation is a behavior of similar or equivalent form that “appears in unconnected, naïve individuals (either in captivity – or in the wild (namely when the [behavior] is also found in culturally unconnected wild populations))” (Bandini & Tennie 2017, p. 7).

The know-how of reinnovatible behaviors (latent solutions) becomes expressed behaviorally via the interaction of individual learning mechanisms with genetic (e.g., possession of the relevant cognitive attributes related to manual coordination) and environmental (e.g., availability of tool raw materials or food items) factors, both of which determine if a behavior *can* be reinnovated and the form that it takes (Bandini & Tennie, 2017; Tennie et al., 2020). In addition to genetics and environment, ontogenetics (e.g., learning during specific parts of the lifetime, *sensitive periods*; Biro et al., 2003; Tennie et al., 2020), non-copying social learning (e.g., Tennie et al., 2009, 2020), stochastic processes, and idiosyncratic factors (e.g., individual personality, motivation levels, social status) can play a significant role in how likely a latent solution is to be reinnovated by any one individual. Expression of know-how can also be influenced indirectly – individually or socially – by the pre-existence of other expressed know-how (pre-existing know-how has the possibility of promoting the development of subsequent know-how, but it might also suppress certain behavioral expressions, e.g., Bernstein-Kurtycz et al., 2020; Heyes, 1994; Marshall-Pescini & Whiten, 2008a).

The full ‘repertoire’ of behavioral forms within the reach of individual reinnovation – including the entire spectrum from those that are grounded more in genetic predispositions to those that are more dependent on (individual) learning processes and are sometimes even socially/culturally influenced (Tennie et al., 2020) – is known as the *zone of latent solutions* (ZLS; Tennie et al., 2009). Contrary to certain misconceptions about the ZLS paradigm (cf. Bandini et al., 2021a; Haidle & Schlaudt, 2020), the ZLS is not related to any “outdated instinct concept” (Bandini et al. 2021a, p. 77) whereby behaviors are essentially genetically predetermined responses to particular stimuli (Bandini & Tennie, 2018; Bandini et al., 2021a). The ZLS is not a completely static element: the number and composition of viable latent solutions can change as a consequence of evolutionary changes to *biological* traits, e.g., cognitive capacities, brain physiology, anatomy, dietary preferences, and niche specialization, not to mention all other factors mentioned above, particularly the environment. This then drives the formation of species-specific ZLS (with only very limited, and even inconsequential, interindividual variation; Reindl et al., 2018; Tennie et al., 2020).

Furthermore, the likelihood by which reinnovation of one or all latent solutions occurs can also vary on the species level (e.g., for nut-cracking; cf. Bandini et al., 2021b; Neadle et al., 2020; Reindl et al., 2016). Neither a single individual nor a single population of a species can be expected to 'actualize' all behavioral forms in their ZLS, because of limitations on time and inherent risks that come with interacting with the environment. The frequency of re-innovation and the varied expression of latent solutions at the individual- and population-level (but not the actual *know-how* of a behavior) is also influenced by non-copying social learning mechanisms (e.g., Bandini & Tennie, 2017; Buskell & Tennie, in press; Tennie et al., 2020).

Minimal cultures

Those cultural processes and outcomes which fulfill the minimal definition of socially-influenced behavior or information are placed in the category of *minimal culture* (Buskell & Tennie, in press; Neadle et al., 2017; Snyder et al., 2022; Tennie et al., 2020). Importantly, this does not refer to one specific process or cultural character and patterning but varies according to species-typical cognitive capacities and the types of information being transmitted and processed (thus, a quick and all-encompassing term for cultural processes that do not particularly involve cultural transmission of know-how).

Non-copying social learning mechanisms play a central role in the generation of minimal cultures, not by influencing the form of behavioral know-how, but rather by increasing the frequency of know-how innovations within a population (e.g., Bandini & Tennie, 2017; Buskell & Tennie, in press; Tennie et al., 2020). These can produce the illusion that know-how innovations are 'spreading' through a population by some means of copying, but instead, these are better described as *socially mediated serial re-innovations* (term first used in Bandini & Tennie, 2017; see also Huffman & Hirata, 2004; Tennie et al., 2010). For example, the presence of fellow group members at a location with termite mounds (know-where) and interacting with sticks (know-what) could increase the likelihood that an individual chimpanzee (*Pan troglodytes*) will visit the termite mounds and manipulate branches and sticks. Given enough time and motivation, this hands-on engagement with the relevant materials – channeled additionally by other biological and cognitive factors – is what leads to the development of know-how (here, for termite dipping) by the individual. Repetition of this process (hence, *serial* re-innovations) is proposed then to explain population-level behavioral patterns that are observed in non-human animals (cf. Acerbi et al., 2022; Robbins et al, 2016; Samuni et al., 2020; van Schaik et al., 2003; Whiten et al., 1999). Though population-scale events are more difficult to observe than ones of the individual level, at least a few examples of such socially mediated re-innovation have been found in wild and captive non-human great apes (hereafter, great apes; Bandini & Tennie, 2020; Hobaiter et al., 2014; Huffman & Hirata, 2004).

Previously, primatologists have made note of cultural patterns (what is often referred to as *traditions*) consisting of differential distribution of cultural traits across populations within the same species (Acerbi et al., 2022). These traditions – known so far from all the great ape species including chimpanzees (Whiten et al., 1999), bonobos (*Pan paniscus*; Samuni et al., 2020), gorillas (*Gorilla gorilla* and *G. beringei*; Robbins et al., 2016), and orangutans (*Pongo pygmaeus*; van Schaik et al., 2003) – have generally been associated with copying social learning mechanisms (e.g., Whiten et al., 1999). This interpretation is often justified by the method of exclusion: in the absence of suitable genetic or ecological explanations, a behavior is considered not just cultural but even likely copied from others (see Whiten et al., 1999; Koops et al., 2015; Stout et al., 2019, for some notable applications of this logic; see also Byrne, 2007; Stout et al., 2010; Stout comment on Tennie et al., 2017). However, observations alone of wild apes cannot directly indicate the mechanisms underlying the acquisition of their behavior (e.g., Bandini et al., 2022; Neadle et al., 2017), the method of exclusion is not exhaustive enough to totally eliminate the possibility of environmental and genetic influences (e.g., Acerbi et al., 2022; Motes-Rodrigo & Tennie, 2021; Neadle et al., 2017; cf. Langergraber et al., 2010; Luncz et al., 2012), and there is a lack of unambiguous evidence for copying abilities in unenculturated and untrained captive great apes (e.g., Clay & Tennie, 2018; Neadle et al., 2021; Tennie et al., 2012).

Recent computational modelling work (Acerbi et al., 2022) reinforces the importance of these caveats about the method of exclusion. Here, we developed an agent-based model, wherein six populations of a fictional species of great ape – named ‘oranzees’ – were placed in a virtual continent, mapped after the six populations of chimpanzees that were studied by Whiten and colleagues (1999). These populations consisted of individuals all latently capable of a broad spectrum of 64 behavioral forms (know-how), which they could potentially re-innovate across their lifetimes. In the model, the likelihood that any individual expresses a behavior is determined by the behaviors that the individual already knows and expresses, the behavioral know-how already expressed by other individuals in the population, and both genetic capacities and ecological affordances (determined by geographic gradients in the modelled continent). By design, none of the oranzees are capable of any sort of copying social learning mechanism, and therefore, cannot be affected by any direct cultural transmission of know-how. Outputs of the model demonstrate that copying-independent pathways (here, socially-mediated re-innovation) can produce population-level distributions of behaviors that are nearly identical to the actual patterns reported by Whiten et al. (1999) for wild chimpanzees. More than that, iterations involving different combinations of values for genetic and ecological factors are able to reproduce this result (in fact, even when genetic and ecological factors are set to zero). At minimum, the potential equifinality (i.e., different processes can produce the same shared outcome; e.g., Acerbi et al., 2016; Barrett, 2019; Perreault, 2019; see also discussion in Whiten et al., 2004 and Davidson & McGrew, 2005) between socially mediated serial re-innovation and form copying means that population-level behavioral distributions are insufficient for determining the mode(s) of cultural transmission in non-humans.

Cumulative culture of know-how

The cumulative cultural evolution of know-how is perhaps the most prominent feature of the human phenomenon (Boyd, 2018; Boyd & Richerson, 1996; Henrich, 2016; Stout & Hecht, 2017; Tennie et al., 2020; Tomasello, 1999). In the past, cumulative culture of know-how and its cognitive underpinnings enabled our ancestors to settle nearly all environments on Earth, to develop otherwise unfathomable technologies, and to create ideas and beliefs that even defy any adaptive reasoning (Boyd, 2018; Henrich, 2016; Kurzban & Barrett, 2012; Stout & Hecht, 2017; Tomasello, 1999). This feature is the driver behind *most* (as human culture can also be minimal or non-cumulative, e.g., Neldner et al., 2020; Reindl et al., 2016, 2022; Snyder et al., 2022; also cf. logic of Donald, 1998) of human cultural diversity, whether that is diversity of material objects or variation of customs, practices, languages, and more (Kurzban & Barrett, 2012; Motes-Rodrigo & Tennie, 2021; Stout & Hecht, 2017; Tomasello, 1999).

This process of cumulative cultural evolution typically involves a ratchet effect (in a sense, a stacking of subsequent or consecutive know-how innovations; Dean et al., 2013; Kurzban & Barrett, 2012; Tennie et al., 2009; Tomasello, 1999, 2003) that ultimately produces behaviors or information that single individuals are incapable of re-innovating from scratch and that thus must be copied to be produced and reproduced (*culture-dependent traits* à la Reindl et al., 2016 or, a more stringent case, *copying-dependent forms* à la Tennie et al., 2020; see also

Boyd & Richerson, 1996; Tomasello, 2003). Additionally, random drift inherent in cumulative cultural evolution of know-how (including basic and unavoidable copying error; Eerkens & Lipo, 2005; for the specific context of knapping and other reductive processes, see also Schillinger et al., 2014) generates a minimum level of radiation or diversification of cultural traits. In order for innovations to ratchet and culture dependent traits to be reproduced and to spread, individuals must have the ability to transmit know-how information via copying (Boyd & Richerson 1985, 1995; Dean et al., 2013; Tennie et al., 2009; Tomasello, 1999; Tomasello et al., 1993; though see for example Stout & Hecht, 2017). Therefore, the mechanistic basis of cumulative culture of know-how is *copying social learning* (Figure 1; see Tennie et al., 2020).

Many authors have promoted the notion that cumulative culture of know-how is not just fundamental to human behavior and cognition, but also that it is a feature unique to humans (e.g., Dean et al., 2012, 2013; Schuppli & van Schaik, 2019; Tomasello, 1999; see also Heyes, 2021). Varied claims of cumulative culture in non-human animals (e.g., Davis et al., 2016; Haidle & Schlaudt, 2020; Saldana et al., 2019; Sasaki & Biro, 2017; Vale et al., 2017) have been put forth, yet these all fail to properly account for observable differences in the kinds and variation of cultural manifestations present in the human species but which are more – if not totally – lacking in non-humans (see Motes-Rodrigo & Tennie, 2021; Tomasello, 1999). Other claims have only posited for copying social learning mechanisms in non-human animals without ‘full-blown’ cumulative culture of know-how (e.g., Byrne, 2009; Byrne & Russon, 1998; Whiten, 2017; Whiten et al., 2003, 2009a), but these claims then lack empirical qualification for the dearth of evidence for ratcheting or copying-error-driven changes in know-how (Eerkens & Lipo, 2005): if humans and non-humans are capable of the same copying social learning mechanisms, then why should only human culture stand out as so starkly ‘inventive’ and diverse (see Boesch & Tomasello, 1998; Dean et al., 2012, 2013; Kurzban & Barrett, 2012; Tomasello 1999)? With regard to the technological domain, there are no primate tool-use behaviors that are unambiguously explainable by copying social learning types or cumulative culture of know-how (see, e.g., Bandini & Tennie, 2020; Dean et al., 2012, 2013; Tennie et al., 2020), although there are claims to the contrary (cf. Byrne, 2009; Byrne & Russon, 1998, Koops et al., 2015, 2022; Marschall-Pescini & Whiten, 2008b; Whiten et al., 2009a). For example, perception of task (behavioral) complexity in toolmaking has been taken as evidence for cognitive complexity, including in relation to cultural cognition and social learning (e.g., Byrne, 2007; Haidle, 2010). Aspects such as

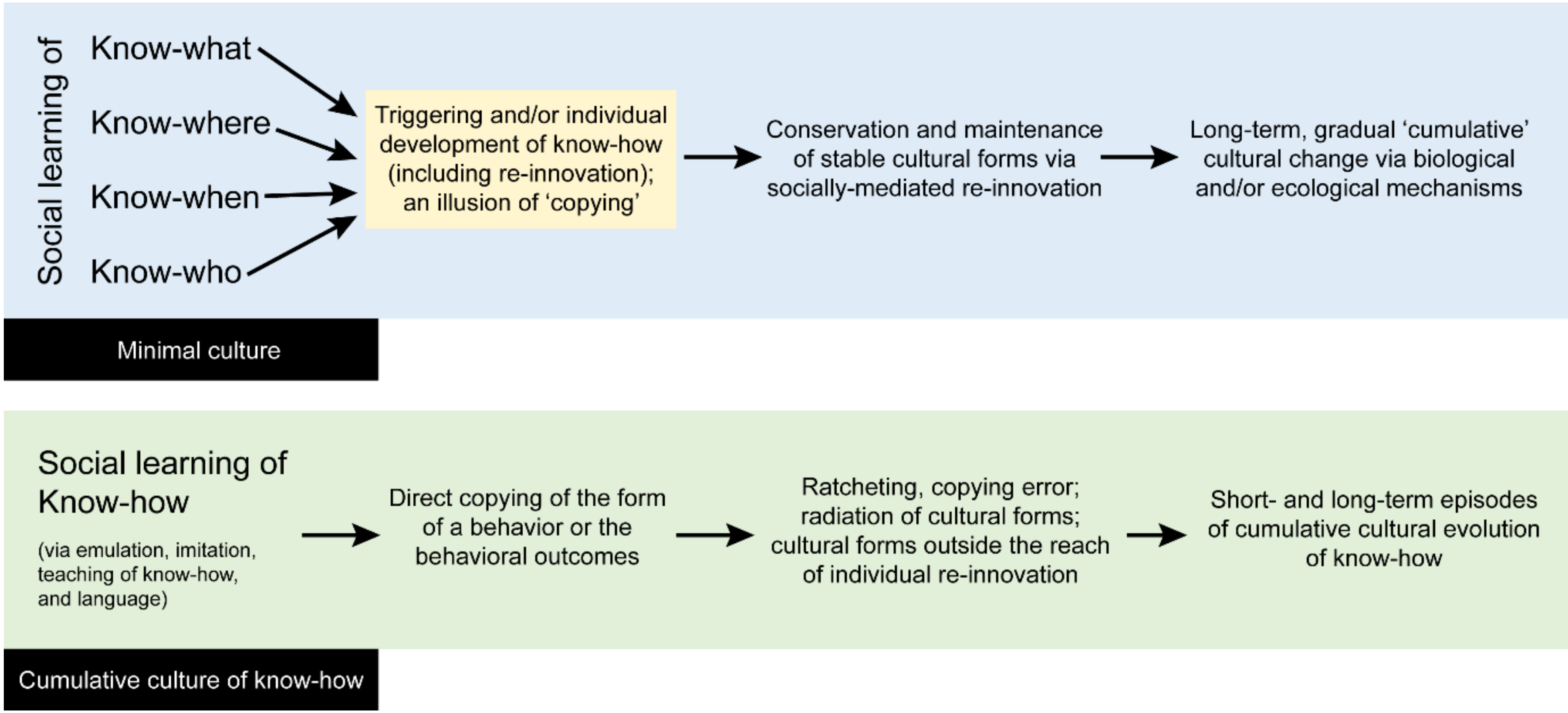


Figure 1 Generalized pathways of cultural transmission and short-term and long-term stabilization of traditions in the categories of minimal culture and cumulative culture of know-how.

combinatorial and compositional tool use (Putt et al., 2022) and action grammars of toolmaking (Stout et al., 2021) have even been studied as a proxy for understanding the possible technological hypothesis for language evolution (e.g., Lombao et al., 2017; Stout & Chaminade, 2012; Stout et al., 2008). These approaches, however, fail to account for the possibility that complex behaviors (with steps and hierarchies; cf. Byrne, 2007) may yet still be guided by relatively simple processes (as in bird nest-building, e.g., Colias & Colias, 1964; see also Bandini et al., 2020 and Tennie et al., 2008).

Insights into the question of non-human cumulative culture are also provided by the study of primate archaeology (e.g., Haslam et al., 2017). 4,300-year-old nut-cracking artifacts from chimpanzees in Côte d'Ivoire show a general character of technological stability (Mercader et al., 2007), though later analyses have suggested some *limited* temporal change in stone tools sizes (Proffitt et al., 2018). A similar archaeological record has been discovered for stone tool use by wild capuchins (*Sapajus libidinosus*) in Brazil, also with modest variation in the food processing traces (Falótico et al., 2019). Though the researchers active at both these sites favor a cultural explanation (in these cases, via imitative/copying processes) for these differences, there is a strong possibility that these cases of non-directional, temporal variation in technology may be the result of ecological changes at each site (see discussion in Proffitt et al., 2018 and Falótico et al., 2019), changes in the resources being exploited by the primates (a demonstrated possibility from the diverse percussive behaviors from wild long-tailed macaques (*Macaca fascicularis*); Proffitt et al., 2021a), or differences in resource exploitation or genetics of the different groups that might have occupied the sites at different times (see Falótico et al., 2019). Nonetheless, primate archaeological sites, even with temporal change in the present technologies, do not show a signature of any directional cumulative cultural evolution nor especially a ratcheting of innovations that would obviously be outside individual capacities (thus, reflecting the same problem of equifinality as with living great ape populations, see above; cf. Acerbi et al., 2022; Robbins et al., 2016; Samuni et al., 2020; van Schaik et al., 2003; Whiten et al., 1999).

Determining mechanisms of transmission

Because the primate archaeological record (cf. Bandini et al., 2022; Falótico et al., 2019; Mercader et al., 2007; Proffitt et al.,2018), naked-eye observation of wild apes (see Bandini et al., 2022), and population-level phenomena in living apes (cf. Acerbi et al., 2022; Robbins et al., 2016; Samuni et al., 2020; van Schaik et al., 2003; Whiten et al., 1999) are ambiguous in their affirmation of any particular mechanisms of social learning and cultural transmission, it is necessary to apply separate approaches that are better suited to disentangling the persistent equifinality problem faced in the study of culture. Here, experimental studies are seen as the gap-filling approach (e.g., Bandini et al., 2020, 2022; Henrich & Tennie, 2017; Tennie et al., 2020), whether that be with unenculturated (i.e., not raised by humans) and untrained captive great apes or other primates or with modern humans (both children and adults, depending on experimental setup).

When regarding the theoretical prospect of distinguishing, e.g., latent solutions from copying-dependent forms, one might consider the following thought experiment (originally proposed by Tomasello, 1999). A child (in his example, a human child) is dropped off on a desert island and left in total isolation. Through some fortuitous means, the child is able to grow up with all of its basic needs met in these otherwise problematic circumstances. As this isolated child develops, those behaviors and means of doing that it produces would be considered within the individual reach for re-innovation (i.e., within its ZLS). Any behaviors that would fail to appear – especially if the necessary raw materials or stimuli are available and the test is repeated with numerous islands and numerous abandoned younglings – could then be considered as candidates for culture-dependent traits (or the stricter concept of copying-dependent forms). From this theoretical vantage, only one case from any one of these ‘island children’ is necessary to serve as proof-of-principle that a behavioral know-how is not culturally- or copying-dependent.

This ‘ideal’ version of the *Island Test* is, of course, problematic for any number of ethical and logistical reasons. Rather than have no experiments and thus only very *vague* information about cultural and cognitive processes (see arguments in Bandini et al., 2022), it is pertinent to maximize the explanatory potential by using stringent experimental control. For instance, one of the key components of island tests for primate candidate cultural behaviors is to investigate naïve (i.e., never performed or been exposed somehow to the target know-how) and unenculturated, untrained primates (e.g., Bandini et al., 2020, 2022; Henrich & Tennie, 2017; Tennie et al., 2020). The latter component is required as rearing by humans (including training to perform specific behaviors, as well as the influences from human modes of communication, social interaction, and artifact forms) can actually enhance the capacities for copying social learning in the affected individuals (e.g., Call & Tomasello, 1996, Furlong et al., 2008; Tomasello & Call, 2004), potentially through documented physiological changes to brain connectivity via enculturation processes (Pope et al., 2018). As such, enculturated primates are highly unrepresentative of the natural cognitive abilities of their wild counterparts. Testing unenculturated apes in captivity allows both for a better capture of species-typical behaviors and cognition and for control of extraneous variables that might interfere with the expression of reinnovatible know-how (not to mention, there is greater certainty about the lifetime experiences of captive individuals than in wild primates, which are more difficult to observe consistently and without interruption).

Island tests (and equivalent baseline testing approaches) have been conducted with regard to a plethora of candidate cultural behaviors in different non-human species (reviewed in Tennie et al., 2020; see also Bandini et al., 2020). Evidence for the possibility of individual re-innovation of know-how has been found for many of these behaviors, including – but not limited to – food-finding behaviors in feral pigeons (*Columba livia domestica*; Lefebvre, 1986) and brown capuchin monkeys (*Sapajus apella*; Dindo et al., 2007); nettle feeding and food cleaning in gorillas (*G. gorilla*; Neadle et al., 2017; Tennie et al., 2008); food washing in representatives of all great ape taxa (Allritz et al., 2013); scooping tool use (Bandini & Tennie, 2017), sponging/sponge-tool use (Kitahara-Frisch & Norikoshi, 1982), and stick-pounding in chimpanzees (Bandini & Tennie, 2019); twig probe use in woodpecker finches (*Cactospiza pallida*; Tebbich et al., 2001); and manufacturing and use of twig tools in New Caledonian crows (*Corvus moneduloides*; Kenward et al., 2005) and, a related species that do not express this know-how in the wild, Hawaiian Crows (*Corvus hawaiiensis*; Rutz et al., 2016). In capuchins alone, an impressive variety of behaviors have been re-innovated in captive testing conditions (e.g., Westergaard & Suomi, 1993, 1994a, 1994b, 1995a, 1995b). Tests of human children have further produced positive evidence for re-innovation capacities, including capacities to replicate not just one but 11 (out of 12) tool-use behaviors known from wild great apes (Reindl et al., 2016). These results were then replicated by cross-cultural comparison of San and Australian children, with a twelfth behavior, nut-cracking, additionally being re-innovated in both populations (Neldner et al., 2020).

Nut-cracking has proven to be one of the behaviors most resistant to re-innovation (see Bandini et al., 2020) and is also worth giving special examination due to the common implication of nut-cracking in models of technological, cultural, and cognitive evolution in the hominin lineage (e.g., Bandini et al., 2022; Bril et al., 2015; Carvalho & McGrew, 2012; Carvalho et al., 2008, 2013; Haidle, 2010; Putt, 2015; Rolian & Carvalho, 2017; Sayers & Lovejoy, 2008; Williams-Hatala et al., 2018). Nut-cracking is one of the more classical tool-use behaviors associated with chimpanzees and has been put forth as one of the main candidates for copying-dependent form among non-humans (e.g., Bandini et al., 2020; Boesch & Tomasello, 1998; Koops et al., 2022; Whiten et al., 1999). Whereas reinnovation did not occur in some studies, i.e., with naïve, unenculturated captive chimpanzees (Neadle et al., 2020) and ostensibly naïve, wild chimpanzees (Koops et al., 2022; it should be noted that the chimpanzees in this study rarely interacted with nuts, therefore potentially lacking the prerequisite motivation to then develop nutcracking behavior; see Tennie & Call, in press), multiple other studies demonstrate that nut-cracking can be reinnovated by other species, including human children (Neldner et al., 2020), orangutans (*Pongo abelli* and *P. pygmaeus*; Bandini et al., 2021b), and capuchins (*Sapajus apella*; Visalberghi, 1987). Upon follow-up investigation (Neadle et al. 2020), chimpanzees were unable to perform a nut-cracking task when provided with ample social learning opportunities (contra supposed "repeated and moderately synchronous matching actions": Whiten et al., 2009, p. 2423; Marschall-Pescini & Whiten, 2008b), suggesting possible suppression of know-how learning due to ecological or ontogenetic (i.e., developmental) factors (compare discussion in Bandini et al., 2021b; Koops et al., 2022; Neadle et al., 2020). Genetic factors also seem to play a determining role in the differential expression of a superficially similar behavior (pound-hammering) in long-tailed macaque subspecies (*Macaca f. fascicularis* versus *M. f. aurea*; Bandini & Tennie, 2018; Gumert et al., 2019).

Besides island tests targeting specific kinds of know-how, other experimental paradigms with naïve, unenculturated primates have also been conducted in order to investigate more generally non-human capacities for copying social learning. Numerous studies (e.g., Clay & Tennie, 2018; Neadle et al., 2021; Tennie et al., 2012) indicate that great apes do not possess the ability to copy *novel* behavioral forms and thus culturally transmit know-how (contra, e.g., Byrne, 2009; Byrne & Russon, 1998; Koops et al., 2015, 2022; Marschall-Pescini & Whiten, 2008b; Whiten et al. 2009a; comments by Gowlett; Luncz & Haslam; Nielsen & Whiten; Shipton on Tennie et al., 2017). With this in mind, it seems rather unlikely (and indeed, not parsimonious[2]) to assume that the origins of cumulative culture of know-how could be found in a common ancestor of humans and the other apes (cf. Schillinger et al., 2015; Shipton, 2010; Whiten 2015, 2016, 2017; Whiten et al. 2009b; see Lucas et al., 2020

[2] Theoretically, one could posit that cumulative culture of know-how was present in a common ancestor of humans but was the capacity for cumulative culture of know-how was then lost in the *Pan* lineage after the split with the *Homo* lineage (see Luncz & Haslam comment on Tennie et al., 2017). This, however, requires the assumption of multiple untraceable evolutionary steps (e.g., as might alternatively be evidenced in a – currently non-existent – pre-LCA material record), and so it is more reasonable to seek for the origins of cumulative culture of know-how somewhere in our own lineage after the divergence with chimpanzees and bonobos.

for claim that verbal teaching pre-dates at minimum the Oldowan; see also Pradhan et al., 2012). The lack of concrete answers from this segment of research suggests that better clues might be found elsewhere: after the divergence of the *Homo* and *Pan* lineages and, thus, somewhere within the hominin lineage and its millions-year-old record of stone tools and other behavioral traces (e.g., Boesch & Tomasello, 1998; Mithen, 1996; Montrey & Shultz, 2020; Tennie et al., 2016, 2017; Toth & Schick, 2018; Tramacere & Moore, 2018; Snyder et al., 2022; Stout & Hecht, 2017; Stout et al., 2019; van Schaik et al., 2019; Wynn & McGrew, 1989; see also Foley & Lahr, 2003).

Early stone tool technologies

Investigation of cognitive and cultural evolution within our lineage and across distinct hominin species presents obvious difficulties. For one, all other hominin taxa are extinct, leaving *Homo sapiens* as the sole representative of a once far more diverse panoply of bipedal apes. Although it is fairly reasonable to assume that hominins possessed material culture that included organic tools (made of, e.g., twigs, branches, leaves, and bones) not unlike those used by living primates (e.g., Ambrose, 2001; Bandini et al., 2022; Gabrić et al., 2021; Haslam et al., 2009; Hovers, 2012; Rolian & Carvalho, 2017; Toth & Schick, 2009), taphonomic processes have – critically – prevented the preservation of non-stone materials in the Pliocene (see possible evidence of bone tools in the early Pleistocene of South Africa: Backwell & d'Errico, 2008; d'Errico & Backwell, 2003). On rare occasions, it is possible to study the endocasts of extinct hominins ("replicas of the inner surface of the bony braincase", Dumoncel et al., 2021, p. 480), but the explanatory potential of endocasts for prehistoric cognition is likely limited (cf. Zollikofer & Ponce de León, 2001; Dumoncel et al., 2021).

Therefore, the comparatively good preservation and the ubiquity of stone tools even in the earliest archaeological record provides researchers a unique opportunity to study human evolution and prehistoric cognition and culture (e.g., Ambrose, 2001; Bandini et al., 2022; Foley & Lahr, 2003; Geribàs et al., 2010; Mithen, 1996; Nielsen, 2012; Schick & Toth, 1993; Stout, 2006; Stout & Chaminade, 2009; Stout & Hecht, 2017; Tennie et al., 2016, 2017; Toth & Schick, 1994, 2018). Though the absence of other behavioral traces poses a likely bias that can influence the inferences and interpretations that cognitive archaeologists make (e.g., Tennie et al., 2016), stone tools nonetheless are a useful proxy for filling in important gaps of knowledge about the origins and evolution of human technological and cultural capacities.

Before the Oldowan industry

Bone surface modifications from Dikika, Ethiopia, (McPherron et al., 2010) and crude fractured stone from West Turkana, Kenya, (Harmand et al., 2015) have been taken as evidence of, respectively, tool use and toolmaking at around 3.3 Ma (therefore contemporaneous with *Australopithecus afarensis* and *Kenyanthropus platyops*; though see criticisms, e.g., Domínguez-Rodrigo & Alcalá, 2016). However, other than these two debated examples, there is a general scarcity of hominin behavioral remains between the divergence of the *Homo* and *Pan* lineage and the sudden appearance of numerous Oldowan stone tool assemblages after 2.6 million years ago (Ma) (Braun et al., 2019; Semaw et al., 1997; Stout et al., 2009; though see recent evidence for habitual tool use from entheseal patterns; Kunze, et al., 2022).

Given interpretation that even early stone toolmaking at sites like Gona, Ethiopia, (Stout & Semaw, 2006) shows sophistication and skill that is beyond that of human novices and trained apes (e.g., Ambrose, 2001; Pradhan et al., 2012; Putt 2015; Stout, 2005; Stout & Hecht, 2017; Stout & Semaw, 2006; Stout et al., 2009; Toth et al., 1993; compare with Whiten et al., 2009b), researchers have pondered over the (more-or-less) absence of pre-Oldowan lithic technologies and often suggested that stone tool use and even stone toolmaking of some kind must have occurred that is intermediate somehow (properties and functions of this hypothetical technology naturally being undetermined) between known ape behaviors and Oldowan knapping (Panger et al., 2002; Putt, 2015; Stout et al., 2009; Whiten et al., 2009b). Generally, it is assumed that pre-Oldowan hominins would have possessed variability in tool use behaviors at least comparable to that of living chimpanzees, with direct comparisons most typically being drawn between Oldowan knapping and chimpanzee nut-cracking behavior (e.g., Ambrose, 2001; Bandini et al., 2022; Bril et al., 2015; Carvalho & McGrew, 2012; Carvalho et al., 2008, 2013; Haidle, 2010; Putt, 2015; Rolian & Carvalho, 2017; Sayers & Lovejoy, 2008).

Supposing that living apes and early hominins possess(-ed) the ability to perform basic flaking activities (see Lewis & Harmand, 2016; Motes-Rodrigo et al., 2021; Putt, 2015; Whiten et al., 2009b), the question then becomes why it did not appear in traceable frequencies until approximately 2.6-2.5 Ma. Indeed, evidence from wild primates demonstrates that they can produce viable cutting edge accidentally during varied percussive activities, and yet, use of cutting tools does not occur (in chimpanzees: Mercader et al., 2002;

Proffitt et al., 2018; in capuchin monkeys: Proffitt et al., 2016; in long-tailed macaques: Proffitt et al., 2022a), suggesting the presence of some impediment to the behavioral innovation of cutting tool use (reviewed in Bandini et al., 2022). Various explanations have been put forth for the seemingly frequent toolmaking of the Oldowan (still only 'occasional tool use', according to the reckoning of Shea, 2017) and infrequent or non-existent condition in the Pre-Oldowan, ranging from evolving cognitive capacities and manual skills (e.g., from the more 'ape-like' state to a hominin one: Ambrose, 2001; de Beaune, 2004; Kivell, 2015; Mercader et al., 2002; Panger et al., 2002; Stout, 2006; Stout & Hecht, 2017; Stout et al., 2009; Wynn & McGrew, 1989; see also Whiten et al., 2009b) to different bodily affordances (use of teeth for accessing the inside of nuts: Bandini et al., 2021b, 2022; Pradhan et al., 2012; Toth & Schick, 2018; Wynn & McGrew, 1989) or changes in ecology and the emergence of a new hominin niche (Foley & Elton, 1998; Foley & Gamble, 2009; Putt, 2015; Toth & Schick, 2011; Whiten et al., 2009b; Wynn et al., 2011).

The Oldowan industry

Whereas the record of toolmaking in earlier periods is more a matter of speculation than measurement, the Oldowan industry boasts a relatively extensive record of artifacts and other remains of hominin activities (e.g., Ambrose, 2001; Shea, 2017; Stout et al., 2009; Toth & Schick, 2018). The first Oldowan assemblages come from sediments from Gona and Ledi Geraru in Ethiopia dating to almost 2.6 Ma (Braun et al., 2019; Semaw et al., 1997). After the initial appearances, Oldowan and similar flaking technologies persist and even re-appear for hundreds of thousands, if not millions of years (Ambrose, 2001; Cueva-Temprana et al., 2022; Foley & Lahr, 2003; Shea, 2017). Oldowan toolmaking was seemingly practiced by several different species of hominins (including members of the genera *Australopithecus*, *Paranthropus*, and early *Homo*), with the identity of the very 'first' toolmakers – if such a thing existed or could even be determined – remaining ambiguous (Ambrose, 2001; Braun et al., 2019; Delagnes & Roche, 2005; Hovers, 2012; Kimbel et al., 1996; Semaw et al., 1997; Shea, 2017; Stout et al., 2019; Tennie et al., 2017; Toth & Schick, 2018).

Oldowan technology, much in the same way as lithic technologies that would later follow, involved the use of percussion – using implements identified by archaeologists as 'hammerstones' and 'anvils' – in order to fracture stone, especially siliceous stone, to generate sharp edges that would be useful for cutting activities (Ambrose, 2001; Cotterell & Kamminga, 1985, 1987; Pelcin, 1997; Roche et al., 2009; Toth, 1985). Based on visually identifiable features of the transformed materials, the specific mechanical process in prehistoric stone toolmaking has been determined to be *conchoidal fracture* (Cotterell & Kamminga, 1985, 1987; Pelcin, 1997), the outcomes of which are differentiable from those of other mechanical processes, such as "uncontrolled split fracture" (Stout et al., 2009, p. 262). The main product of conchoidal fracture is the detachment of pieces or stone chips, known as *flakes*, which can be identified on the basis of (qualitative) traits, including a bulb of percussion (or bulb of force), platform (the surface that was struck during the formational percussive event), and ripples and undulations, which show the path followed by the wave of force that passed through the material while fracturing (Cotterell & Kamminga, 1985, 1987; Pelcin, 1997). Thus, conchoidal fracture in the context of archaeological stone toolmaking is also often referred to simply as *flaking*[3] (e.g., Bril et al., 2010; Nonaka et al., 2010; Stout & Semaw, 2006; Stout et al., 2009). The stone objects from which flakes are removed (i.e., objects that were or are being flaked) are usually referred to as *cores*, while prior to any removal event, these would be considered as *blanks*.

As a consequence of repeated flaking events, whether by single individuals or groups of hominins (see Ferguson, 2003; Stout & Semaw, 2006; similarly for the transport of raw materials, see McGrew et al., 2019; Reeves et al., 2021), assemblages were formed, consisting of cores and flakes, as well as other detached pieces and percussive elements. Said predominance of cores and (unretouched, whole) flakes (Shea, 2017) in Oldowan and Mode 1 assemblages has led to them being labelled as 'simple' core and flake technologies or industries (Toth, 1985; Toth & Schick, 2018). Some researchers have assigned later Oldowan (from the early Pleistocene of Africa) assemblages to the Developed Oldowan (Leakey, 1971, 1975), citing a higher proportion of spheroid artifacts and retouching of flakes (see Ludwig, 1999; Toth & Schick, 2018), while others have offered an alternative classification wherein earlier sites (before approximately 1.9 Ma) are considered as Pre-Oldowan (de Lumley et al., 2009; Shea, 2017). Like Isaac (1976), we subsume these into the broader category of Oldowan *sensu latu*, as the technological principles (i.e., the behavior and know-how

[3] Contra Bril et al. (2010), we do not draw a strong distinction concerning intentionality between 'flaking', which they define only as "the production of flakes", and 'knapping', which they associate with "the removal of a succession of flakes to produce an end product of a specific shape" (p. 826). The definition of knapping implemented here refers to any of the percussive behaviors related to prehistoric stone toolmaking ('flaking' and 'knapping' essentially being flipsides of the same coin, focusing respectively on the mechanical and behavioral aspects of the same action or action sequences). Also, contra Putt (2015), we do not limit our definition of 'knapping' to only the freehand technique, but include also, e.g., passive hammer, bipolar, and projectile techniques under the umbrella term of 'knapping'.

involved) remain more-or-less intact, while the differences in the artifact types might yet have non-cultural explanations (cf. Hiscock, 2015; Hovers, 2012; Toth & Schick, 2018).

Reduction sequences from the Oldowan are said to have involved ‘least-effort’ flaking, meaning that the entire focus of reduction was the efficient or expedient production of usable cutting flakes (Isaac, 1984; Key et al., 2020; Moore, 2020; Moore & Perston, 2016; Toth, 1985; Toth & Schick, 2018). Toolmaking was opportunistic, representing the creation of

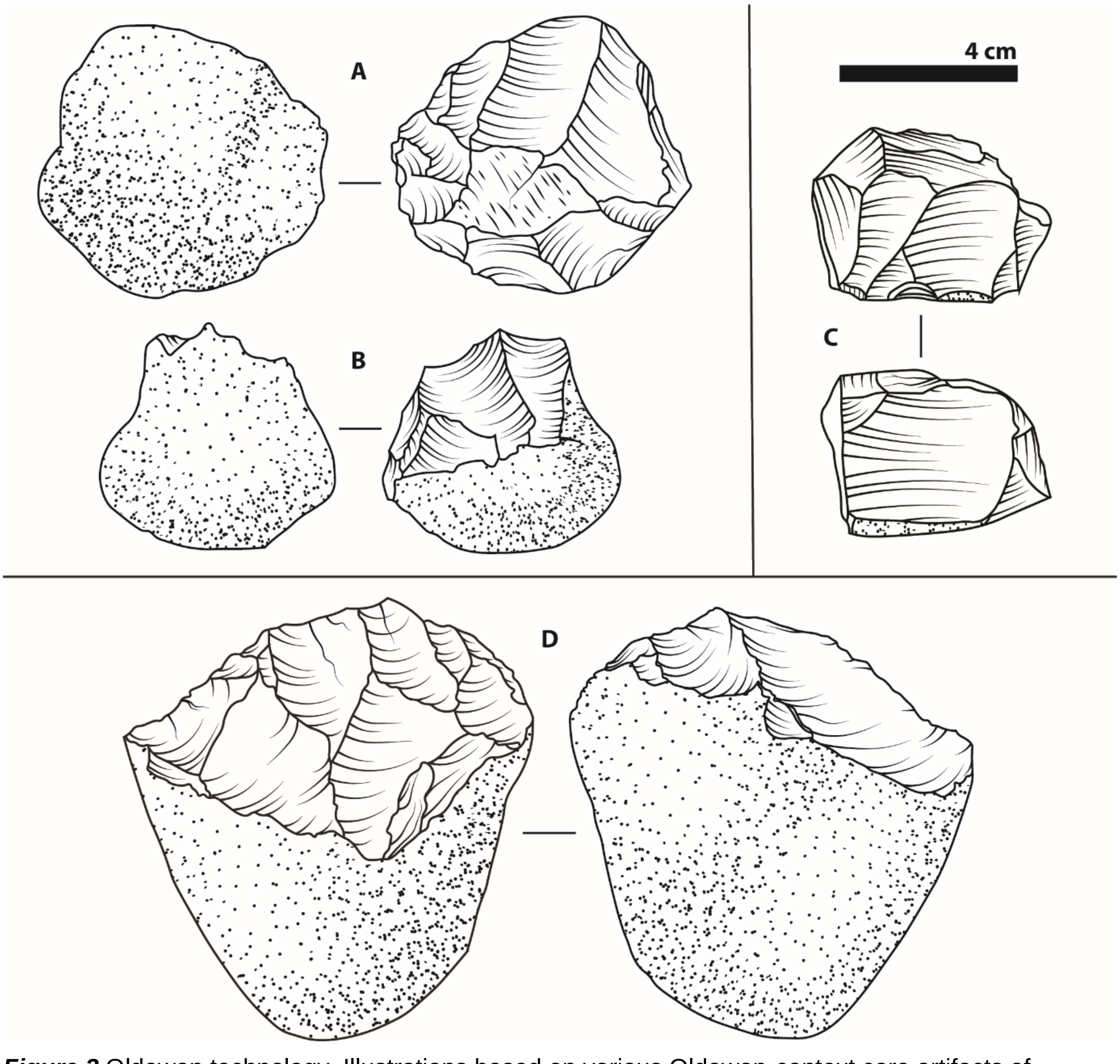


Figure 2 Oldowan technology. Illustrations based on various Oldowan-context core artifacts of differing reduction sequences and raw materials. The oldest, with an age of approximately 2.6 to 2.5 million years (Ma), are from Gona, Ethiopia (Semaw, 2006) and include a radial core (**A**) and a unifacial chopper (**B**). An example of a multifacial core (**C**) comes from the site of Melka Kunture, Ethiopia, where premodern hominins exploited obsidian (volcanic glass) as a raw material, starting around 1.7 Ma (Piperno et al., 2009). The final example is a bifacial core (**D**) from Hadar, Ethiopia and dated to approximately 2.3 Ma (Kimbel et al., 1996). In accordance with archaeological convention, solid lines are used for the outlines of scars (where flakes have been removed from a core surface), curved, incomplete lines represent the ripples from conchoidal fracturing of the knappable material, and dotted areas represent cortex (the unaltered, original surface of the raw material).

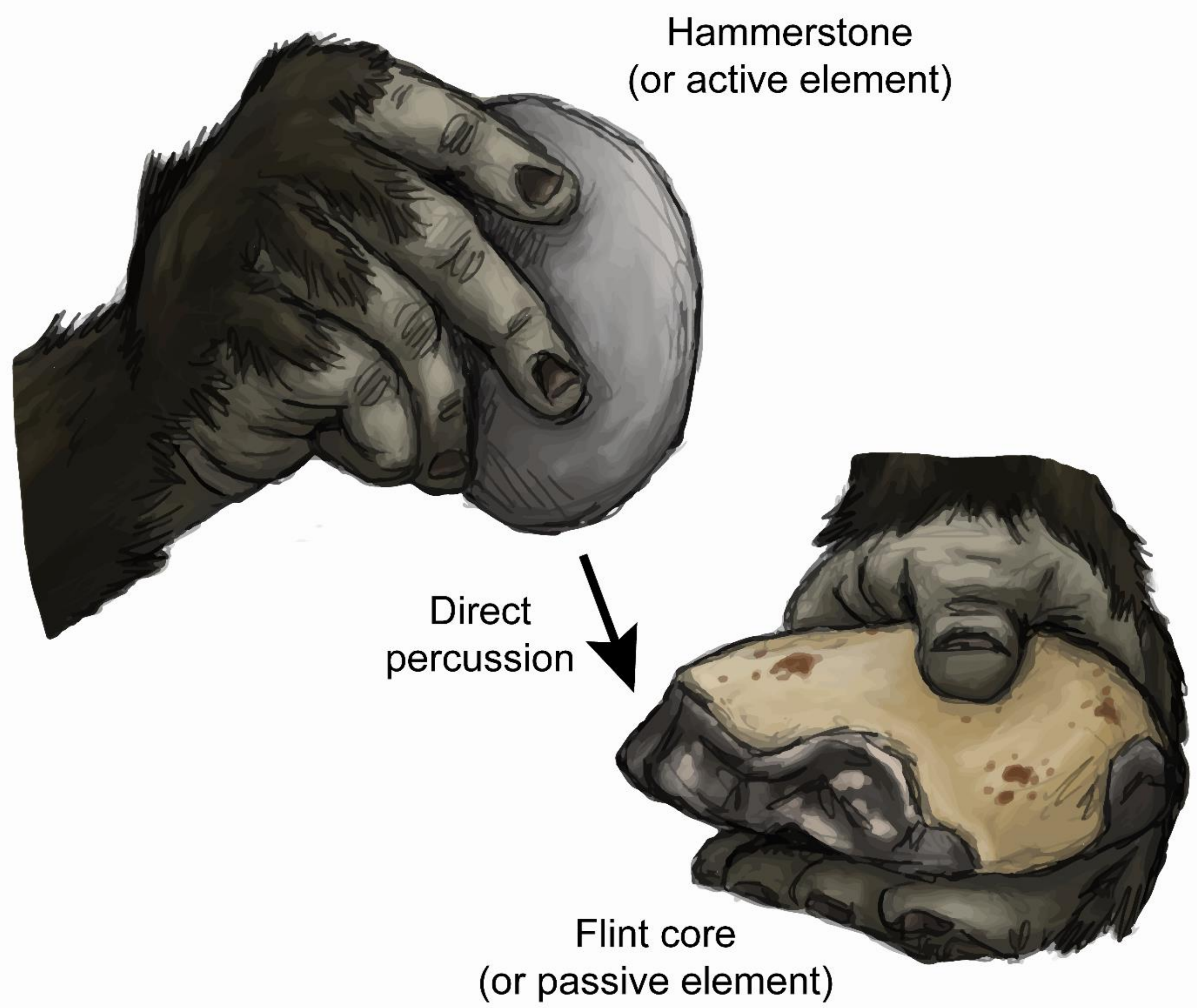


Figure 3 Example of freehand knapping technique, whereby a flint core (a passive element) in the left hand (of a hominin) is struck via direct percussion with a hammerstone (an active element) in the right hand.

usable implements from adjacent raw materials (cf. Key et al., 2020) or mobile toolkits from more distant raw materials (see example discussions of raw material procurement and transport in Duke et al., 2021; Harmand, 2009; Hovers, 2012; McGrew et al., 2019; Reeves et al., 2021; Stout et al., 2005, 2009; Toth & Schick, 2011, 2018) when they were needed (Putt, 2015; Toth, 1985). The resultant artifact types in the Oldowan are interpreted to be unintentional by-products (Figure 2; Moore, 2020; Moore & Perston, 2016; Toth, 1985) rather than the consequence of some sort of mental template or intentional shaping process as ostensibly appeared with the following Acheulean industry (e.g., Ambrose, 2001; Lycett & Gowlett, 2008; Shipton, 2010; Whiten et al., 2003; comment by Wynn on Tennie et al., 2017).

The traditional behavioral paradigm associated with Oldowan technology includes not just the production of cutting edge in the form of flakes or on cores (Figure 3), but also the subsequent usage of viable cutting tools (again, a major departure from the unintentional flaking in other primate species: Mercader et al., 2002; Proffitt et al., 2016, 2018, 2022a; see also the study of Motes-Rodrigo et al., 2022 wherein toolmaking and tool use were separate and non-sequential) for extractive foraging and other tasks, whether that be butchering for flesh consumption, the acquisition of marrow from animal long bones, or various means of processing vegetation for food or secondary toolmaking (Ambrose, 2001; Guerbuez & Lycett, 2021; Hayden, 2015; Hovers, 2012; Mora & de la Torre, 2005; Toth, 1985; Toth & Schick, 2018; Williams-Hatala et al., 2018; Wynn & McGrew, 1989; see also Domínguez-Rodrigo, 2009).

Though the status of the Oldowan as a single, homogeneous industry or cultural entity is oft-debated (see Barsky, 2009; Delagnes & Roche, 2005; Shea, 2017; Toth & Schick, 2011, 2018), artifacts and assemblages from the Pliocene and early Pleistocene of Africa (Oldowan *sensu stricto*; see Foley & Lahr, 2003; Toth & Schick, 2018) share many – if not all – of the same technological features (temporal and cross-ecological stasis: consider the recent work of Cueva-Temprana et al., 2022). This remains true even when one expands comparisons to include artifacts and assemblages from the early to late Pleistocene (sometimes even Holocene) of Europe and Asia, which are defined, alternatively, as either Oldowan *sensu latu* or Mode 1 (following Grahame Clark's classification system of technological modes: see Clark, 1969; Foley & Lahr, 2003; Shea, 2013; Toth & Schick, 2018). Inclusion of non-African 'simple' stone tools also expands the pantheon of hominin taxa (including *H. antecessor*, *H. floresiensis*, *H. erectus*, Middle Pleistocene or Neandertal lineage hominins, and modern humans) that are associated with the technology (see Barsky, 2009; Carbonell et al., 1999; Mithen, 1996; Gabunia et al., 2000; Lombao et al., 2022; Moore & Brumm, 2009; Tennie et al., 2017). Indeed, Mode 1 artifacts were produced by nearly all toolmaking hominins (Foley & Lahr, 2003). Hereafter, the term Oldowan is not applied to infer any strict, unified culture (or tradition), but to the Oldowan *sensu stricto*, while the term Mode 1 is applied to all similar technologies from outside Africa.

The culture(s) of Oldowan hominins

Evidence of anthropogenic objects is often equated with evidence of modern human-like intentions – or at minimum, some kind of intentionality – and culture (Baber & Janulis, 2020; Bamforth & Finlay, 2008; Mithen, 1996; Renfrew, 2004; Stout, 2006; Stout et al., 2015), and as such, the discovery of deeply archaic, early stone tools in eastern Africa has been and is viewed as an indication of modern human-like behavior, cognition, and/or culture in pre-modern hominins even over 2 million years ago (Lombao et al., 2017; Stout et al., 2010, 2019; Whiten, 2015, 2016).

Even though Oldowan and Mode 1 technologies are typically associated with fairly basic technical goals (Isaac, 1984; Moore, 2020; Moore & Perston, 2016; Toth, 1985) and required know-how, which was simple and easy-to-acquire in comparison to other later technologies (Ambrose, 2001; Apel, 2001; Harlacker, 2003; Hiscock, 2015; Muller et al., 2022; Stout & Hecht, 2017; Stout et al., 2015, 2021), many archaeologists associate them with mechanisms of cultural transmission of know-how (i.e., copying). Indeed, the origins (or at least presence) of actualized cumulative culture of know-how have even been directly pinpointed to the earliest Oldowan (e.g., Lombao et al., 2017; Stout et al., 2019) and cultural transmission of know-how is assumed (axiomatically) to have been necessary for the existence of widespread and numerous Oldowan assemblages across Africa and through time (cf. Caruana et al., 2013; Lombao et al., 2017; T. Morgan et al., 2015; Schick et al., 1999; Shipton & Nielsen, 2015; Stout & Semaw, 2006; Stout et al., 2010, 2019; Toth et al., 1993; Whiten, 2015).

Claims for cultural transmission of know-how

A wide range of claims have been made with regard to the supposed necessity of copying social learning mechanisms to perform either the basic act of flaking or the specific techniques and gestures used in Oldowan knapping sequences, including:

- via unspecified mechanism (Duke & Pargeter, 2015 regarding bipolar technique; Eren et al., 2020; Shipton, 2020; Stout & Chaminade, 2007; Stout & Semaw, 2006; Stout et al., 2010; comment by Stout on Tennie et al., 2017),
- via emulation (T. Morgan et al., 2015; Nielsen & Whiten comment on Tennie et al., 2017; Toth & Schick, 2018),
- via imitation (Caruana et al., 2013; Gabora & Steel, 2020; Lombao et al., 2017; T. Morgan et al., 2015; Nishiaki, 2019; Schick & Toth, 1993, 2018; Schick et al., 1999; Shilton, 2019; Shipton, 2010; Shipton & Nielsen, 2015; Stout & Khreisheh, 2015; Stout et al., 2011, 2019; Toth et al., 1993; Whiten 2015),
- via (proto-)language of any kind (Stout & Khreisheh, 2015 regarding “relatively simple, Lower Paleolithic, knapping skills”, p. 869; Lombao et al., 2017 regarding the ‘alternating method’; Cataldo et al., 2018),

- via gestural communication (Cataldo et al. 2018; Gärdenfors & Högberg, 2017; "mimetic communication" a la Shilton, 2019, p. 131; compare also with Putt et al., 2014 for the Acheulean and Ohnuma et al., 1997 for Levallois),
- via verbal communication (Lucas et al., 2020; Lombao et al., 2017 regarding the 'alternating method'; but see comment by T. Morgan on Tennie et al., 2017),
- and teaching (B. Morgan et al., 2015 regarding bipolar knapping; Stout & Khreisheh, 2015 regarding "relatively simple, Lower Paleolithic, knapping skills", p. 869; Lombao et al., 2017 regarding the 'alternating method'; Gabora & Steel, 2020; Lucas et al., 2020; Torres & Preysler, 2020).

Though not all of these authors have directly implicated cumulative culture of know-how, they still characterize Oldowan toolmaking as dependent upon the very cultural transmission mechanisms that enable cumulative culture of know-how and that lead to the evolution of know-how outside of the individual reach for innovation (Boyd & Richerson, 1985, 1995; Dean et al., 2013; Tomasello, 1999; Tomasello et al., 1993; Tennie et al., 2009). The argumentative bases for these claims are varied and occasionally, contradictory, when comparing the views of different authors that advocate for cultural transmission of know-how in the Oldowan.

Although some researchers have explicitly characterized Oldowan stone toolmaking as being – in one way or another – 'ape-like' (more in line with, say, chimpanzee nut-cracking than subsequent lithic technologies) rather than 'human-like' (e.g., Davidson & Noble, 1993; Haidle, 2010; Pradhan et al., 2012; Putt et al., 2017; Stout, 2005; Whiten et al., 2003; Wynn & McGrew, 1989; Wynn et al., 2011; see also comments by Luncz & Haslam and Morgan on Tennie et al., 2017 for criticism of human versus ape cognition dichotomy, along with a claim for cultural transmission in apes and extinct hominins), these comparisons still assume for shared human and ape – and therefore also early hominin – capacities for cultural transmission of know-how (e.g., "Different Oldowan localities exhibit different *socially learned* knapping *procedures*"; Wynn et al., 2011, p. 194, emphasis added by the present author) that are very typical of modern human cognition but that are only dubiously known from apes (Bandini & Tennie, 2020; Clay & Tennie, 2018; Dean et al., 2012, 2013; Neadle et al., 2021; Tennie et al., 2012, 2020). Others share the *a priori* assumption that hominins possessed copying abilities because apes supposedly possess copying abilities (e.g., Shipton, 2010; Whiten, 2015, 2016; Whiten et al., 2009b; comments by Gowlett, Luncz & Haslam, Nielsen & Whiten, and de la Torre on Tennie et al., 2017), but then press that Oldowan behavior and culture might be more complex than in apes (therefore, ‘higher-fidelity’ copying mechanisms) or might require the transfer to a different ecological niche (see, e.g., Shipton, 2010; Stout & Hecht, 2017; Whiten et al. 2009b; see also Pradhan et al., 2012).

Artifacts from many Oldowan site, including some of the very oldest Oldowan artifacts (Stout & Semaw, 2006; Stout et al., 2009), show indications of skillful control of the mechanical fracture of stone (Ambrose, 2001; Braun et al., 2019; Delagnes & Roche, 2005; Duke et al., 2021; Rein et al., 2014; Stout, 2005; Stout & Semaw, 2006; Stout et al., 2009, 2010, 2019). The process of knapping is regarded – by at least some authors – as difficult (e.g., Eren et al., 2020; Shipton, 2020; Sterelny & Hiscock, in press; comments by Shipton and de la Torre on Tennie et al., 2017) and therefore costly to acquire (including in terms of time and raw materials; e.g., Ferguson, 2003; Muller et al., 2022; Pargeter et al., 2019, 2020; Shelley, 1990; Stout, 2002) so difficult and costly (also dangerous) as to require copying social learning (Hiscock, 2015; Pargeter et al., 2019, 2020; Sterelny, 2021; Sterelny & Hiscock, in press) in order to overcome the potential risks and the lost time and effort involved in skill acquisition (Lycett et al., 2015; Schillinger et al., 2015; comment by Lycett on Tennie et al., 2017). Since the toolmaking skill is perceived to be beneficial, or even necessary, to fitness and survival (Pargeter & Shea, 2019; Pargeter et al., 2019, 2020; Shea, 2017, 2019; Toth & Schick, 2018), it has also been suggested that expedient learning would be required and that expedient learning would be inherently social (i.e., in the sense that is advantageous for living modern humans to learn by copying others). Thus, evidence that Oldowan hominins were proficient at the task of knapping is inferred to also be evidence for cultural transmission of know-how (e.g., comment by Shipton on Tennie et al., 2017; see Sterelny & Hiscock, in press).

Ostensible empirical support for this comes from knapping experiments with novice human volunteers (e.g., Cataldo et al., 2018; Lombao et al., 2017; T. Morgan et al., 2015; Pargeter et al., 2021). In these studies – where, in all cases, the participants were provided ample opportunities for copying social learning, whether that involved the display of ‘finished’ artifact forms, action demonstrations, direct teaching, or some other know-how information – there was a correlation between these opportunities to learn (including with respect to the specific learning mechanisms and contexts) and the quality of the participants’ knapping performances (e.g., teaching and language were demonstrably for efficient at transmitting knapping-related know-how than imitation and emulation: T. Morgan et al., 2015). On the basis of said results, cultural transmission of know-how – of one type or another – was concluded to not only be necessary for learning Oldowan knapping in the present human condition, but also was extrapolated to be necessary for the acquisition of toolmaking skills by extinct hominins during the Oldowan (Cataldo et al., 2018; Lombao et al., 2017; T. Morgan et al., 2015; Pargeter et al., 2021). Extending the 'just-so' logic that individual learning could not possibly be sufficient for adopting stone toolmaking skills (see above), these studies also *a priori* consider emulation to be the minimal possible learning condition (for them, it is the control condition: see T. Morgan et al., 2015), as emulation involves "merely seeing finished products (as Oldowan knappers would, of course, have done)" (comment by Nielsen & Whiten on Tennie et al., 2017, p. 660).

Finally, regional patterning and site-to-site differences in knapping behaviors have also been taken as evidence of cultural transmission of know-how (e.g., Barsky, 2009; Delagnes & Roche, 2005; de la Torre, 2019; Hiscock, 2015; Stout et al., 2010, 2019; Toth & Schick, 2011, 2018), often utilizing similar logic to – and facing similar epistemological problems – inferences of cultural transmission mechanisms from population-level behavioral patterns in living primates (cf. Acerbi et al., 2022; Robbins et al., 2016; Samuni et al., 2020; van Schaik et al., 2003; Whiten et al., 1999). Additionally, the identification of proposed localized traditions of Oldowan knapping behaviors stand in contrast to the variable stasis across time and space that has been noted by other observers (Cueva-Temprana et al., 2022; Foley & Lahr, 2003; Gallotti, 2018; Isaac, 1972, 1984; Jelinek, 1977; Semaw et al., 2003; Tennie et al., 2016, 2017; van Schaik et al., 2019); technological stasis in both apes and early hominins is yet viewed by some as unproblematic for interpretations of copying social learning as the extensive temporal and geographic conservatism in know-how can supposedly be a product of highly faithful maintenance of transmission and conformity (Gowlett, 1996; Lycett et al., 2015; T. Morgan et al., 2015; Shipton, 2010; Tramacere & Moore, 2018; Whiten, 2017; see also Foley & Lahr, 2003, p. 119).

An alternative hypothesis

Although the notion is certainly popular among cognitive archaeologists and technologists nowadays – especially provided the current mainstream trends in primatology and experimental archaeology (see above) – the assumption for cultural transmission of know-how is not universally shared among all researched who have published on the matter of Lower Paleolithic cognition (e.g., Corbey, 2020; Corbey et al., 2016; Cueva-Temprana et al., 2022; Davidson & McGrew, 2005; Horta et al., 2022; Hovers, 2012; Mithen, 1996; Montrey

& Schultz, 2020; Nielsen, 2012; Richerson & Boyd, 2005; Tennie et al., 2016, 2017; comments by Davidson and Wynn on Tennie et al., 2017; van Schaik et al., 2019; see also account by Nonaka et al., 2010).

The stasis in technological know-how throughout the Oldowan constitutes one of the main arguments against the presence of cultural transmission of know-how among early toolmakers (e.g., Cueva-Temprana et al., 2022; Hovers, 2012; Montrey & Schultz, 2020; Tennie et al., 2016, 2017; van Schaik et al., 2019). Even more so, the repeated re-appearance of 'simple' flaking strategies, not just within the Oldowan, but also elsewhere in the world in later periods has been taken as evidence that the technological principles were not entirely so difficult for individual hominins to grasp (Apel, 2001) and then become fixed at the group-level (e.g., Cueva-Temprana et al., 2022; Braun et al., 2019; Hovers, 2012; Shea, 2017; Tennie et al., 2017). According to both proponents and opponents of inferences for cultural transmission of know-how in the Oldowan, the limited differences in artifactual and behavioral forms between Oldowan sites separated by time and/or space could be explained by factors such as the biology and cognition of the various implicated hominin taxa, ecology, foraging purposes and requirements, and properties of the available raw materials (Ambrose, 2001; Delagnes & Roche, 2005; Domínguez-Rodrigo, 2009; Duke et al., 2021; Foley & Lahr, 2003; Goldman-Neuman & Hovers, 2011; Hovers, 2012; Nielsen, 2012; Proffitt et al., 2021b; Semaw et al., 1997; Shea, 2017; Stout & Semaw, 2006; Stout et al., 2009; Tennie et al., 2016, 2017; Toth & Schick, 2011, 2018; Toth et al., 1993; Wynn & McGrew, 1989; cf. Stout et al., 2019).

It is rather evident that previous experimental studies (e.g., Cataldo et al., 2018; Lombao et al., 2017; T. Morgan et al., 2015; Pargeter et al., 2021) have demonstrated a correlation between opportunities for special types of copying social learning and improved performance at knapping tasks, but none of these previous studies have undertaken a true control condition. Instead, these have viewed emulation or 'reverse engineering', a mechanism capable of transmitting know-how (Caldwell & Millen, 2009; Reindl et al., 2017), as the minimally effective possible learning condition for Oldowan hominins (see T. Morgan et al., 2015; comment by Nielsen & Whiten on Tennie et al., 2017). These prior approaches, therefore, completely leave out the possibility that individual re-innovation could be a sufficient process for the generation of knapping-related know-how in naïve persons. As such, previous research has placed excessive emphasis on the best or most efficient conditions for *modern humans* to adopt early knapping techniques (here, this includes passive hammer, bipolar, freehand, and projectile techniques: de la Torre, 2019; Harmand et al., 2015; Putt, 2015; Toth & Schick, 2011, 2018; Wynn & McGrew, 1989), failing to acknowledge that potentially none of these copying social learning mechanisms and learning situations were available to any or all Mode 1-practicing hominins prior to the origin of our species. By this very essence, prior work in experimental archaeology has failed to demonstrate what could be considered the minimum *necessary* competences for cognition in the Oldowan.

Furthermore, knapping experiments generally share an additional problem: that of ecological validity. Whereas Oldowan hominins are believed that have produced flakes as tools when needed for miscellaneous foraging tasks, human participants in knapping studies are usually required to reduce cores in single sessions from start to finish. This means that experimental participants are trained to knap in contexts completely divorced from the actual purpose of toolmaking: tool use (note that there are a handful of exceptions to this pattern in human knapping experiments whereby the participants were asked to test the utility of their created tools: Bisson, 2001; Stout & Semaw, 2006; Stout et al., 2009).

According to an alternative hypothesis (i.e., one that does not implicate the cultural transmission of know-how in the proliferation of early stone tool technologies), the know-how of early knapping techniques was within the reach of individual re-innovation or, in other terms, within the ZLS of our hominin progenitors (Tennie et al., 2016, 2017). As with some earlier accounts of the cultural and cognitive capacities of Oldowan (or Mode 1) hominins (see Davidson & McGrew, 2005; Hovers, 2012; Mithen, 1996; Nielsen, 2012; Nonaka et al., 2010; see also perspective presented in comment by Davidson on Tennie et al., 2017), non-copying social learning, such as local and stimulus enhancement, played a role in the cultural transmission of other information types like know-where and know-what (Montrey & Schultz, 2020; Tennie et al., 2016, 2017; van Schaik et al., 2019). This non-know-how pathway for cultural transmission would lead to triggering of know-how development by subsequent individuals and the stabilization of minimal cultures (i.e., socially mediated serial re-innovations; Bandini & Tennie, 2017; Buskell & Tennie, in press; Montrey & Schultz, 2020; Tennie et al., 2010, 2020).

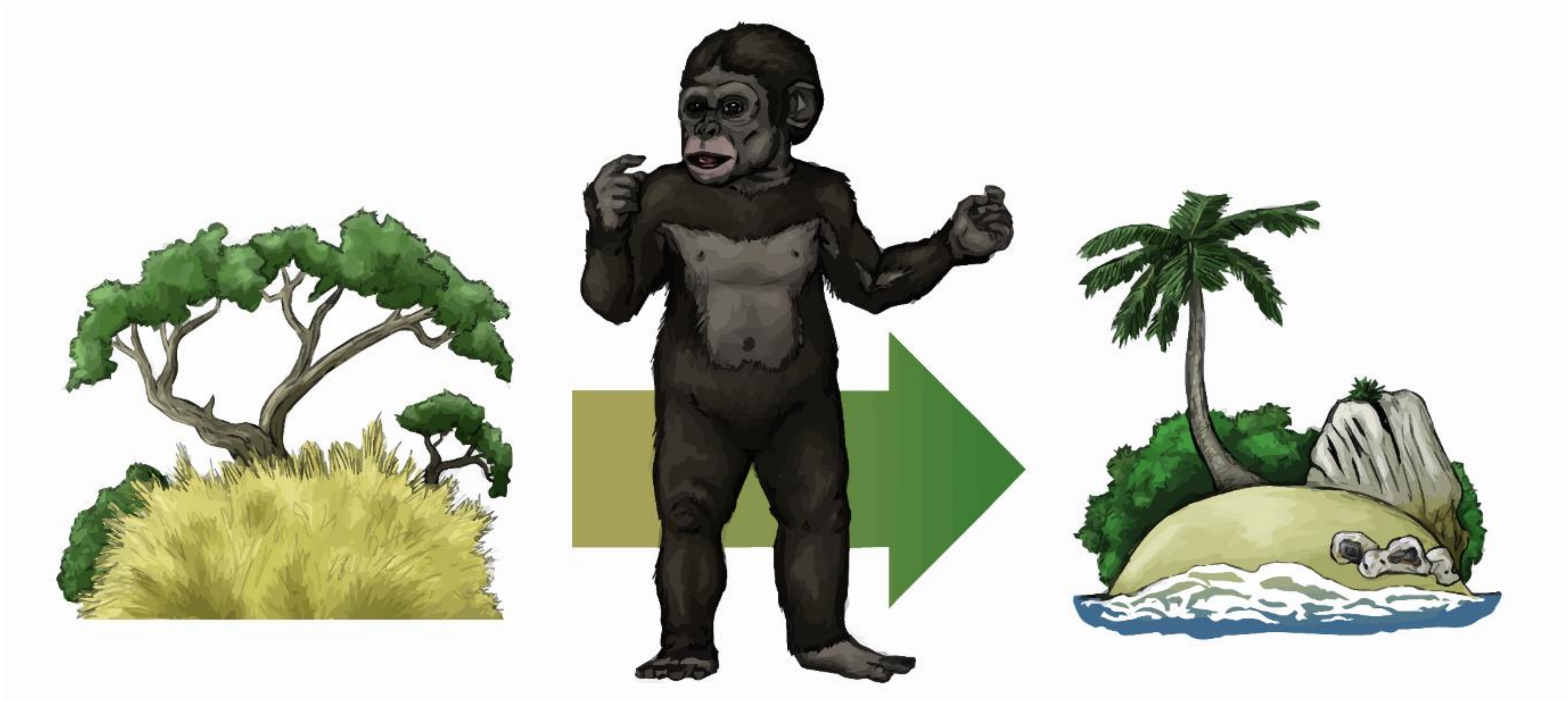


Figure 4 In an idealized version of the island test (were one to have a time machine in order to perform such a test), a young hominin is taken from its native environment and placed in isolation (i.e., alone on an island). For the thought experiment, all needs of the hominin are met (e.g., no possibility of starvation, no detrimental effects of social isolation; see Tomasello, 1999; Tennie et al., 2017). The island that it is left to inhabit has all materials that would allow for and ecological contexts that would motivate toolmaking. The purpose of the test would then be to wait and see if the individual would eventually re-innovate any or all early knapping techniques.

Testing the alternative hypothesis

Though detractors of the ZLS hypothesis for early stone tools have challenged the testability of the hypothesis (see Shipton, 2020; comments by Luncz & Haslam, Shipton, and Stout on Tennie et al., 2017), the hypothesis is indeed built on an empirical approach to the determination of re-innovative potential of human and non-human behaviors: the Island Test (Figure 4; Tennie et al., 2016; Tomasello, 1999). As previously described, real-life island tests are controlled experiments in which naïve individuals, provided with pertinent raw materials and motivation, are tested for their ability to re-innovate the know-how of target behaviors (Bandini et al., 2020; Tennie et al., 2020).

Tests of the knapping-related re-innovative capacities of naïve primates have already been conducted (of capuchin monkeys: Westergaard & Suomi, 1994a, 1995a; of chimpanzees: Bandini et al., 2021c; of orangutans: Motes-Rodrigo et al., 2022), in each case motivating the primates with a puzzle box baited with food (serving as a simulation of extractive foraging: Wright, 1972). The results of these studies were mixed. Chimpanzees, at least in the relevant study (Bandini et al., 2021c), proved incapable of either producing or using cutting tools (perhaps mirroring the lack of re-innovation in nut-cracking tests: Koops et al., 2022; Neadle et al., 2020). Orangutans were shown capable of spontaneously re-innovating one of the relevant knapping techniques, passive hammer technique, but without subsequent use of the created tool (Motes-Rodrigo et al., 2022). Spontaneous *use* of a cutting tool only appeared in a separate testing condition. Of these primates, capuchin monkeys were – despite greater phylogenetic distance from humans and hominins compared to the studied ape – actually the most starkly (re-)innovative (Westergaard & Suomi, 1994a, 1995a). Naïve capuchin monkeys were able to re-innovate freehand, passive hammer, bipolar, and projectile techniques for flake production, as well as the use of flakes to access a puzzle box award (Westergaard & Suomi, 1995a). This result is perhaps not so surprising (though with hindsight), given the documentation of unintentional flaking events in wild capuchin monkeys (Proffitt et al., 2016).

Despite the certainly informative outcomes of naïve primate knapping experiments, the most phylogenetically relevant (see Bandini et al., 2022; Snyder et al., 2022) species has not yet been tested for its baseline knapping abilities: living modern humans. The central focus of this doctoral project and dissertation is experimental tests – modelled on the protocols from ape experiments (Bandini et al., 2021c; Motes-Rodrigo et al., 2022) – of human participants kept in cultural isolation (i.e., contra previous studies, not presented with any opportunities for the cultural transmission of know-how, such as demonstrations, emulation models, or teaching). We tested individuals recruited from the local area (N = 28; n_{male} = 14, n_{female} = 14) for their ‘problem-solving’ skills: could they potentially reproduce the know-how of any or all early knapping techniques, as well as the artifactual outcomes of ‘simple’ knapping solutions, with only the motivation of a baited puzzle box (a reward of money instead of food, in these experiments) and without even the term ‘stone tool’ ever being mentioned during testing? Additionally, all participants were tested in isolation (to prevent transmission of any know-how between and across participants, and also to serve as independent replications of the experimental outcomes) during four-hour sessions (to hopefully provide enough time for re-innovations to occur), and evaluation of their naivety before the beginning of the session was conducted post-test with a questionnaire, as a way of avoiding that a questionnaire on stone tool experiences before testing might cue them in or channel them towards the behaviors of interest.

We hypothesized that naïve modern humans would be capable of re-innovating the principal knapping techniques (know-how) involved in Oldowan toolmaking and that the products of this toolmaking would be similar to artifacts from the known archaeological record. Initiation of conchoidal fracture to produce flakes by any or all of these techniques in modern humans would be taken as proof-of-principle for independent re-innovation of knapping know-how in extinct hominin species. At the same time, this would warrant a re-appraisal of hypothesis for cultural transmission of know-how in the Oldowan as being less parsimonious. Indeed, the results of our study show that naïve individuals are capable of all early knapping techniques and of producing Oldowan-like knapping outcomes (Snyder et al., 2022).

In the remainder of this dissertation, I will elaborate upon this introduction and brief summary of the main experiments with the following outline of content:

- **Chapter II**: a review of the experimentation that has been conducted with novice knappers, along with key considerations for epistemological problems in cognitive archaeology and potential ways forward despite said problems,
- **Chapter III**: an overview of our tests of potentially standardizable raw materials, offering a unique comparison of the techniques and materials that can be used in controlled knapping experiments,
- **Chapter IV**: the materials and methods used in the final study (Snyder et al., 2022) along with various obstacles and difficulties involved in research with human volunteers,
- **Chapter V**: the main results of our study of naïve knappers,
- **Chapter VI**: a discussion of the study on naïve knappers,
- And **Chapter VII**: a general summary of the methodological, empirical, and theoretical contributes to research of the cognition and cultural capacities of our hominin forebearers.

In the appendices, there is supplementary text related to the methods and results of the experiment, as well as the documents that were used during the experimental procedure.

II. Novice knapping experiments: possibilities, pitfalls, and potential in cognitive archaeology

Explanation of candidate contribution to collaborative work

This chapter is based on collaborative work. The text represents the current version of the following manuscript:

Snyder, W.D. & Tennie, C. (in prep). Novice knapping experiments: possibilities, pitfalls, and potential in cognitive archaeology.

Conceptualization: C.T., W.D.S. Investigation: W.D.S. Visualization: W.D.S. Formal analysis: W.D.S., C.T. Writing (original draft): W.D.S. Writing (review and editing): W.D.S., C.T.

Empirical studies, in the form of controlled knapping experiments with novice human subjects, have become a central component of research into the evolution of human and hominin behavior and cognition over the course of the last several million years. These empirical studies have ostensibly provided insights into the processes and mechanisms (e.g., social learning mechanisms) that are involved in the production of prehistorical stone tool technologies. This line of research relies on the building of inferences and analogies that bridge the gap between processes present in modern experimental settings and the end-products that we can find in the preserved archaeological record. Inference building presents an array of epistemological problems concerning the validity of assumptions made about past behavior and cognition. Nonetheless, we argue that novice knapping experiments can still be a valuable data resource, so long as researchers responsibly acknowledge the assumptions implicit in their protocols and models.

Introduction

The replication of prehistoric stone technologies has been a part of archaeological praxis since at least the 1860s, when pioneering researchers first applied their personal observations of and experiences with gunflint preparation to recreate artifacts from the archaeological record (Johnson, 1978). Experimental knappers of the 19th and 20th centuries were interested in demonstrating that stone tools were indeed anthropogenic objects, in unveiling the manufacturing process of these tools, and in studying fracture mechanics and other physical properties of different material types. Today, stone artifact replication serves not only to aid our understanding of how these objects likely would have been produced (and the types of gestures, techniques, and physical processes involved: e.g., Bril et al., 2010; Parry et al., 2014; Nishiaki, 2019; Rein et al., 2013, 2014), but also can teach us about the cognition and behavior of the very organisms who produced them (Schick & Toth, 1993; Stout et al., 2009); Toth & Schick, 1994; this is further informed by a synthesis of data from field archaeology, primatology, and ethnographic fieldwork and field experiments (e.g., Bril et al., 2005; Flenniken, 1984; Roux et al., 1995; Stout, 2002). Finally, archaeologists view – as informed by personal experience, ethnography, and analogical reasoning – the knapping of stone to produce usable tools as a skill, being honed through practical experience and attention (e.g., Bamforth & Finlay, 2008; Stout & Hecht, 2017; Wynn, 1991). In recent years, this latter aspect, variation in skill and learning, has taken center stage in experimental archaeological research (cf. Stout et al., 2009).

Presently, the status of stone knapping as a *copied* skill (see copied know-how; Bandini et al., 2020; Snyder et al., 2022; Tennie et al., 2020) and the proposed manner(s) of transmission are indeed under debate, with claims ranging between stone tools being evidence of some form of language (e.g., Cataldo et al., 2018; Lombao et al., 2017; Stout & Khreisheh, 2015) or pedagogy (e.g., Stout et al., 2014; Tehrani & Riede, 2008; Torres & Preysler, 2020) on one side to (certain) stone tool types being a product of genetic predetermination on the other (Corbey, 2020; Corbey et al., 2015) – as well as some more intermediate positions (see Snyder et al., 2022; Tennie et al., 2016, 2017). Regardless of the implied (individual or social learning) mechanisms, we can generally assume that hominins were not born precocially ready to start knapping away at stones (cf. Shelley, 1990) and most likely underwent a period of naivety and relative inexperience (i.e., novicehood).

Although in previous times there were any number of hominin species that would have been engaged in stone toolmaking behaviors (e.g., Semaw et al., 1997; Stout et al., 2019; Toth & Schick, 2018), there is now only a single species left from the lineage: *Homo sapiens*. By this token, we study living modern humans (e.g., Schick & Toth, 1993, 1994) and also living non-human primates (hereafter, primates: cf. Bandini et al., 2022) as model organisms (i.e., stand-ins for extinct hominin species) in knapping experiments. Despite the potential issues with the construction of such analogies combined with questions of validity about the contexts of toolmaking in experiments versus the past, knapping experiments are a pivotal component to mapping out hypotheses and models for the evolution of stone tool technologies and their makers. The insights to be gained from said experimentation are certainly beyond what one can understand about these evolutionary processes solely by observation and measurement of artifacts from the record (Binford, 1972; Killin & Pain, 2022).

In the present review, we aim to provide an overview of knapping studies conducted with human novices. Here, we consider studies relevant if the participants fit the *sensu latu* definition of a novice (i.e., an individual with little or no experience with stone knapping), if the experiments use explicitly knappable materials (thereby excluding studies that involve replacement tasks: e.g., Lycett et al., 2015; Rolian et al., 2011; Schillinger et al., 2014), and if they are reported in the peer-reviewed literature (barring a few exceptions that have major contributions to the discussions in this developing research field). These studies have produced a plethora of new information relevant to our understanding of human evolution and both small- and large-scale patterns in the archaeological record, but at the same time, they are susceptible to theoretical and methodological shortcomings. From our perspective, overcoming these issues will require the examination of all reasonable methodologies and avenues for interpretation, while staying cognizant of issues of experimental control, internal and external validity, and the priors and biases of both researchers and test participants (e.g., Eren et al., 2016; Lin et al., 2018).

Novice knapping experiments

Who is a novice?

Flintknapping experiments that include or focus on novices are a growing contingent within cognitive archaeology. Early examples of directed experimentation with novice knappers involved local excavation assistants at Koobi Fora (Toth, 1982). The workers were asked only to produce flakes from selected cobbles without being provided examples of Oldowan cores. According to Toth (1982), these workers could flake stone, producing Oldowan-like unifacial cores (replicated also by American university students, Toth, 1985), and with additional toolmaking practice, the assistants attempted bipolar, passive hammer and free-hand percussion techniques. From this, Toth concluded that "Koobi Fora hominids… appear to have understood [bifacial flaking as the most efficient way to reduce a cobble]" while novice flakers "do not initially understand this principle" (Toth, 1982, p. 248). In this seminal work, however, Toth did not provide any further details about the excavation assistants beyond the fact that they were roughly familiar with the raw materials and the artifact types due to their involvement in excavating the site and that they had never knapped before.

Very much as the term would imply, a *novice* in the context of this review is any individual who has relatively little or no experience with knapping or stone tools (cf. Ferguson, 2003; Stout, 2006; Stout & Chaminade, 2007; Stout & Semaw, 2006; Stout et al., 2009; Toth, 1982, 1985). The term novice is often used interchangeably with similar terms, such as 'non-expert' or 'apprentice' (see e.g., Lombao et al., 2017); for our purposes, we particularly reject the term 'apprentice' as this might lead to confounding associations with the kinds of instruction and social systems, which are typical of apprenticeships in human societies (e.g., of the Langda adze-makers; Stout, 2002) but otherwise cannot be assumed for extinct hominin populations.

In identifying someone as a novice, two kinds of experience are typically highlighted: hands-on experience with knapping (i.e., savoire-faire, procedural knowledge, or practical know-how; Bamforth & Finlay, 2008; Pargeter et al., 2020) and conceptual knowledge of stone tools (i.e., connaissance; Bamforth & Finlay, 2008; Pargeter et al., 2020). Individuals who have no direct, hands-on knapping experience can be regarded as being ‘technique-naïve’ (e.g., Snyder et al., 2022), while those who have neither practical experience nor any conceptual knowledge of stone tools can be regarded as ‘totally naïve’ (see Snyder et al., 2022; contra definition of ‘complete novice’ in Geribàs et al., 2010; Stout & Semaw, 2006). In recruiting novices for knapping studies, various criteria have been applied. In some cases, participants have been expected to have simply no hands-on, knapping experienced (e.g., Bisson, 2001; Ferguson, 2003; Lombao et al., 2017; Putt, 2015; Putt et al., 2014, 2017; Stout, 2006; Stout & Chaminade, 2007; Stout & Semaw, 2006; Stout et al., 2009, 2011, 2014). In other studies, qualification as a novice was determined by a maximum limit of pre-study knapping experience (often set at a number of hours or years; see Nonaka et al., 2010; Stahl, 2008; Sternke & Sørensen, 2007; Stout et al., 2014; Williams et al., 2010; Zorrilla-Revilla et al., 2021) or was determined by evaluation of participants’ knapping products and allotment to skill groups (see Bril et al., 2015; Parry et al., 2014). Only very rarely have researchers reported the pre-study experiences of participants with stone tools beyond the hands-on or practical dimension (e.g., going so far as to investigate the explicit context of information types and exposure to stone tools; only one study has verifiably studied truly totally-naïve novices, see Snyder et al., 2022).

Most novices in stone-knapping studies have come from WEIRD backgrounds (Henrich et al., 2010), with a notable overrepresentation of university students and affiliated academics from nations such as the United States, United Kingdom, France, and Japan (see Figure 5). In only a handful of rare cases (in controlled studies) are the participants are from non-WEIRD populations (e.g., Stout & Semaw, 2006; Toth, 1982, 1985). As far as age, most fall within their 20s, most likely corresponding to the high recruitment of college-age individuals. Experimental novices tend to lack the familial-social contexts and ecological prerogatives associated by many authors with the stone knapping behaviors (e.g., Pargeter et al., 2019, 2021; Stout & Khreisheh, 2015). Several studies have tried to replace this artificially by recruiting participants specifically for their interest in the study and motivation to learn how to knap (Pargeter et al., 2019, 2020; Stout & Hecht, 2015, 2017; Stout & Khreisheh, 2015).

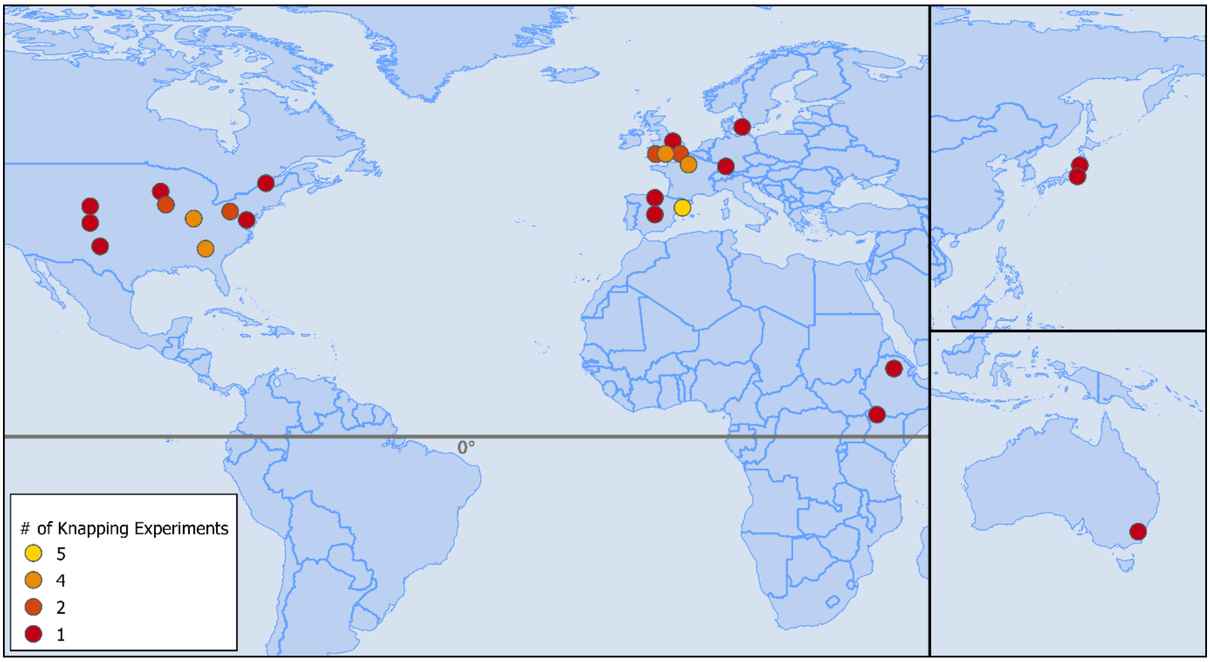


Figure 5 **G**lobal distribution of approximate locations where novice knapping experiments have been performed. It is generally assumed that most participants in these studies come from the countries where the studies were performed or other WEIRD countries. In rare cases, countries of origin are reported that differ from the testing location (e.g., participants from Mexico, Argentina, and Iraq in Bargallo & Mosquera, 2014). Map by Patrick Cuthbertson.

Why do we test novices?

Given the assertion that hominins were not born with fully attuned knapping skills (e.g., Shelley, 1990), knapping is understood as a behavior or set of behaviors that would have been learned over an extended period and perhaps even with identifiable phases of skill development (cf. Pargeter et al., 2019, 2020; Stout et al., 2009). As such, "much of what is cognitively interesting about stone knapping is likely to occur in the process of learning rather than during expert performance" (Stout, 2006, p. 283). Human novices are presumed to be a better model for the hominin learners than experts would be, e.g., in terms of the degree of sensorimotor control and coordination and the suite of gestures used for knapping (cf. Bisson, 2001; Putt, 2015; Stout & Semaw, 2006). The inverse is also argued to be true, with novices used as a control group to which expert human knappers and 'habitual' toolmakers from prehistory (see Marzke et al., 1998). Studying novices as opposed to expert knappers also provides further benefits, such as serving as a more cognitively and culturally diverse population (e.g., allowing for research of handedness effects on toolmaking, since there are relatively few left-handed expert knappers; Bargallo & Mosquera, 2014; Bargallo et al., 2018). Novices are also – suspected to be – less susceptible to biases due to prior knowledge, which might otherwise skew and confound the results of behavioral and cognitive research (i.e., the “knapper’s conceit”; Eren et al., 2016; see also logic of Bisson, 2001; Marzke & Shackley, 1987; Putt, 2015; comment by Davidson on Tennie et al., 2017). Not only do novice knapping studies have important implications for studying human cognition and the evolution thereof, but they also might also contribute to our understanding of artifactual variation in stone tools assemblages and the record as a whole (Ferguson, 2003; Proffitt et al., 2021b; Shelley, 1990; Stout & Semaw, 2006; Stout et al., 2009; Torres & Preysler, 2020).

Novice knapping studies have intended to investigate a fairly wide scope of research themes from archaeology and evolutionary anthropology, including those related to intentionality of stone tool morphologies (Bisson, 2001; Shipton & Clarkson, 2015; Toth, 1982, 1985), the identification and characterization of skill and skill levels (Bril et al., 2010; Duke & Pargeter, 2015; Ferguson, 2003; Geribàs et al., 2010; B. Morgan et al., 2015; Nonaka et al., 2010; Pargeter et al., 2019, 2020; Shelley, 1990; Sternke & Sørensen, 2007; Stout, 2005; Stout & Semaw, 2006; Stout et al., 2009, 2014; Toth, 1982, 1985; Zorrilla-Revilla et al., 2021), cognition and cultural transmission (Cataldo et al., 2018; Harlacker, 2003; Lombao et al., 2017; T. Morgan et al., 2015; Nonaka et al., 2010; Pargeter et al., 2019, 2020, 2021; Putt et al., 2014, 2017; Ohnuma et al., 1997; Shipton, 2020; Stahl, 2008), handedness of toolmakers (Bargallo & Mosquera, 2014; Bargallo et al., 2018), visuomotor processes during skill acquisition (Bayani et al., 2021), cognitive neuroscience of stone toolmaking (Gabrić et al., 2021; Hecht et al., 2015; Putt et al., 2017, 2019; Stout, 2005, 2006; Stout & Chaminade, 2007; Stout et al., 2011, 2015), the technological precursor(s) of knapping (Bril et al., 2015; Putt, 2015), and the biomechanics of toolmaking, including gestures, grips, muscle recruitment, and limb movements (Marzke et al., 1998; Nishiaki, 2019; Parry et al., 2014; Rein et al., 2013, 2014; Williams et al., 2010; Williams-Hatala et al., 2018, 2020)

There is a focus on Early and Middle Stone Age technologies, certainly in recent knapping studies related to cultural transmission (e.g., Cataldo et al., 2018; Lombao et al., 2017; T. Morgan et al., 2015; Pargeter et al., 2019, 2020, 2021; Putt et al., 2014, 2017), but generally, there is a wide range regarding the stone technologies being utilized or examined in these studies; the total range of examined technologies includes *Pan* nut-cracking techniques (Bril et al., 2015) up to and including technologies from Mesolithic Europe and Paleoindian North America (e.g., Ferguson, 2003; Sternke & Sørensen, 2007). If nut-cracking is interpreted as a behavior shown by the last common ancestor (LCA) of chimpanzees and humans (e.g., Bandini et al., 2022; Carvalho & McGrew, 2012; see also Koops et al., 2010), there is also a very large temporal representation of technologies, spanning at least five million years. The combination of both temporal and spatial dispersion (the geographic distribution spans continents) of experimentally tested technologies renders broad and finite inter-study nigh impossible.

How do we test novices?

Skill studies can be either cross-sectional or longitudinal – though a mixed approach is theoretically possible (Stout et al. 2009). Cross-sectional studies draw conclusions about skill development by comparing separate individuals of differing skill levels – a between group approach (e.g., novice versus expert; cf. Bril et al., 2015; Duke & Pargeter, 2015; Marzke et al., 1998; Sternke & Sørensen, 2007; Stout & Semaw, 2006; Stout et al., 2009, 2011, 2014). Cross-sectional studies have the advantage of not requiring intensive, long-term training investments during the study; on the other hand, these experiments cannot control for intergroup differences arising from confounds like individualized preferences, abilities, and differing learning contexts (see e.g., Putt et al., 2017; Rein et al., 2014; Stahl, 2008). Cross-sectional studies with short training periods have also been used to compared the effects of different learning conditions on skill development and learning (e.g., Cataldo et al., 2018; Lombao et al., 2017; T. Morgan et al., 2015; Pargeter et al., 2021; Putt et al., 2014, 2017; Ohnuma et al., 1997). Longitudinal studies are within-group studies, which focus on skill development of individuals over time, providing more explicit data about the skill acquisition process itself (rather than about outcomes only) - as well as allowing for the comparison of individual learning curves – e.g., before versus after (cf., Pargeter et al., 2019, 2020; Stout et al., 2009). The disadvantage of this approach is the need for time investment and commitment from both experimenters and participants. As a result, longitudinal studies (especially long-lasting ones) are rare, perhaps even non-existent (see Pargeter et al. 2019, 2020, 2021). Slight variations of both paradigms have been adopted in the literature. Given a general interest in the cultural qualities of knapping skills, so-called transmission chains, an approach borrowed from the field of cultural evolution studies, have also been applied (where subject *A* learns first, then subject *B* learns *A*, and so on; this design was used in T. Morgan et al., 2015).

Though the exact learning conditions in prehistory are still unknown (e.g., Putt et al., 2017) and learning conditions in controlled knapping experiments may not be valid representations for that precise unknown quantity (see below), said learning conditions are still thought to

Table 2 Types of information that participants in novice knapping studies are exposed to just prior to or during testing. ✓ indicates that information types are reported as present in the relevant references, while X indicates that information types are specifically excluded.

Study	Formal archaeological education (students or professional archaeologists)	Info sheet on stone tools or knapping	Briefing on concepts of stone tools or knapping	Verbal request to produce flakes or hit material with hammer	Verbal request to perform specific knapping techniques, gestures, actions	Verbal request to test viability of cutting tools	Emulation model (i.e., desired artifact form)	Imitation model (i.e., visual demonstration of knapping)	Imitation model (i.e., group sessions with multiple observable knappers)	Basic teaching or scaffolding	Gestural teaching	Verbal teaching
[1]	✓							✓	✓		✓ (NV)	✓ (V)
[2]	✓		✓	✓		✓	✓					
[3]	✓ (Some)		✓	✓				✓				
[4]				✓		✓	✓					
[5]	✓ (Some)	✓	✓					✓	✓			
[6]	✓				✓			✓		✓	✓	✓
[7]							✓	✓	✓		✓ (NV)	✓ (V)
[8]				✓	✓			✓ (FH)				
[9]		✓				(?)	✓ (RE)	✓ (I/E)		✓ (BT)	✓ (NV)	✓ (V)
[10]					✓			✓			✓ (NV)	✓ (V)
[11]								✓			✓ (NV)	✓ (V)
[12]		✓		✓			✓	✓ (- C)			✓ (G, CT)	✓ (SO, CT)
[13]			✓	✓			✓	✓	✓		✓	✓
[14]			✓	✓			✓	✓	✓		✓ (UT)	✓ (UT)
[15]	X		✓					✓	✓		✓ (NV, V)	✓ (V)
[16]	X	X	X	X	X	X	X	X	X	X	X	X

NV = nonverbal condition, V = verbal condition, FH = freehand knapping condition, - C = all except for control condition, G = gestural teaching condition, CT = complete teaching condition, SO = speech-only teaching condition, ? = unclear if present or not, RE = reverse engineering condition, I/E = imitation/emulation condition; BT = basic teaching condition, UT = unrestricted teaching condition

[1] Ohnuma et al., 1997; [2] Bisson, 2001; [3] Harlacker, 2003; [4] Stout & Semaw, 2006; Stout & Chaminade, 2007; Stout et al., 2009; [5] Stout et al., 2011, 2014; [6] Hecht et al., 2014; Stout et al., 2015; [7] Putt et al., 2014; [8] Putt, 2015; [9] T. Morgan et al., 2015; [10] Lombao et al., 2017; [11] Putt et al., 2017; [12] Cataldo et al., 2018; [13] Pargeter et al., 2019, 2020; [14] Pargeter et al., 2021; [15] Gabrić et al.,

represent reasonably situations for skill acquisition that can be used for inference-making about the past. A wide variety of conditions have been applied in knapping experiments. As mentioned above, some studies utilize two or more differing learning contexts, in order to build up comparisons between the contexts and make evolutionary interpretations from those comparisons. Some very typical conditions for knapping experiments related to social learning include emulation conditions (i.e., involving the provisioning of model artifacts; e.g., T. Morgan et al., 2015), imitation conditions (i.e., involving action demonstrations; e.g., Harlacker, 2003; T. Morgan et al., 2015; Stout et al., 2011, 2014), and nonverbal and verbal teaching conditions (e.g., Cataldo et al, 2018; Lombao et al., 2017; Putt et al., 2014, 2017; Ohnuma et al., 1997). The latter pairing of conditions have become particularly popular with respect to the proposed relevance for studying the origins and evolution of language capacities in the human lineage (first suggested by Toth & Schick, 1994 and first applied by Ohnuma et al., 1997). In nearly all studies, these learning conditions involve the participants being imparted with social information (i.e., from the experimenters and/or fellow participants) related to just stone tools or stone tools and toolmaking (see Table 2). Even using terms like 'core' or 'flake' can have the potential to influence – in some way – behavioral outcomes and learning trajectories of participants (see Putt et al., 2017; Snyder et al., 2022). As such, relatively little attention has been paid to true baseline conditions for human acquisition of knapping skills (cf. Tennie et al., 2016, 2017); some studies (e.g., Cataldo et al., 2018; T. Morgan et al., 2015) have claimed to have baseline or control conditions (all of which minimally included emulation opportunities), but actual baselines should completely exclude transmission of social information related to tools or toolmaking know-how (as was pursued in Snyder et al., 2022).

What have we learned from novices?

Perhaps the first step toward understanding the skill acquisition process – before even attempting external validation – is to record how modern human novices learn. Experimentalists have long had thoughts into this phenomenon (e.g., observations of erroneous biface-making technique in Schick, 1994; Shelley, 1990), and controlled experiments have offered an excellent means for studying these developments more empirically in modern humans. These experiments can tell us, for example, what sort of behaviors and mistakes are typical of modern novices (especially as opposed to other skill

Table 3 *Observed markers of novice toolmaking, including typical behaviors and corresponding artifact attributes, which theoretically would be traceable in the archaeological record.*

Novice behaviors	Artifact attributes
More frequent and consistent errors [1], [2]	
Less frequent correction of errors and infrequent removal of errors using 'pick-up flakes' [1], [3]	
Higher rate of unsuccessful percussion attempts (e.g., strikes without removals, misplaced blows, less accurate percussion) [1], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16] Selection of blanks with natural imperfections [1]	Signs of battering [1], [2], [11], [12], [16], [17], [18], [19]
Striking of the core at inconsistent distances from the core edge or striking more often at the periphery of the core [18], [19]	Thinner platforms [18], [19]
Lack of capacity for invasive flaking of a core surface [12], [16], [19], [20], [21]	Smaller flakes [11], [12], [16], [19], [22], [23] Smaller flakes produced more easily than larger flakes [5], [6], [11], [12] Thinner flakes [16], [18]
Striking of the core further away from the core edge [10], [24]	Thicker platforms [10], [24]
Less likely to use pre-existing scars/removals as knapping platforms [16], [22]	Proportionally higher representation of unifacial cores [16], [22]
More likely to strike core in approximately the same place (i.e., on the same platform) [1], [22]	More dorsal scars on flakes [22] Stacked step terminations [1], [17]
Less uniform or inconsistent knapping technique, opportunistic selection of available platforms [10], [18]	
Use of too little force when striking cores [1], [9], [12]	Multiple incipient fractures, more flake fragments and shatter [1], [2]
Use of excessive force or less control over application of force [5], [11], [13], [19]	More step and hinge terminations and fewer feather terminations [1], [2], [18], [22]
Use of longer movements [5], [11], [19]	
More likely to resort to anvil-based techniques (e.g., bipolar) when control over flaking is limited [2], [17], [23], [25]	
Less intensive reduction or exhaustion of cores [1], [16], [18], [22], [23], [26]	Higher rates of artifacts that show no signs of removals [1], [9], [16], [18], [22]
	Fewer flakes produced [16], [22]
Less successful at and require more attempts for splitting cobbles using axial bipolar technique [7]	
More frequent core discards due to inability to remove additional flakes or elimination of suitable platforms for further knapping [1], [16]	Fewer artifacts that are recognizable as traditional 'tool types' [10], [16], [22]

[1] Shelley, 1990; [2] Sternke & Sorensen, 2009; [3] Muto, 1974; [4] Bayani et al., 2021; [5] Bril et al., 2010; [6] Bril et al., 2015; [7] Duke & Pargeter, 2015; [8] Lombao et al., 2017; [9] B. Morgan et al., 2015; [10] Nishiaki, 2019; [11] Nonaka et al., 2010; [12] Pargeter et al., 2020; [13] Parry et al., 2014; [14] Rein et al., 2013; [15] Rein et al., 2014; [16] Stout et al., 2009; [17] Ferguson, 2003; [18] Stout & Semaw, 2006; [19] Zorrilla-Revilla et al., 2021; [20] Stahl, 2008; [21] Stout et al., 2014; [22] Harlacker, 2003; [23] Pargeter et al., 2021; [24] Gabrić et al., 2021; [25] Geribàs et al., 2010; [26] Lombao et al., 2017

levels). Generally, novice mistakes (see Table 3) are interpreted from a lack of understanding (i.e., know-how) related to the functional parameters of knapping and a lack of sensorimotor coordination attained only after prolonged practice (e.g., Bril et a., 2010; Nonaka et al., 2010; Pargeter et al., 2020; Rein et al., 2013, 2014; Zorrilla-Revilla et al., 2021). The 'errors' made by novices include step or hinge terminations (e.g., Shelley, 1990; Sternke & Sørensen, 2007; Stout & Semaw, 2006), unsuccessful removals and battering (e.g., Nonaka et al., 2020; Pargeter et al., 2020; Shelley, 1990; Sternke & Sørensen, 2007), and small, thin flakes from non-invasive removals on cores (e.g., Pargeter et al., 2020; Stahl, 2008; Stout et al., 2009, 2014; Zorrilla-Revilla et al., 2021). However, expert knappers can also produce knapping errors (e.g., Torres & Preysler, 2020), but do so at lower rates than novices and also being more capable of 'correcting' errors, e.g., with the removal of so-called 'pick-up' flakes (Muto, 1974; Shelley, 1990). A further caveat regards raw material: the traits associated with novice knapping actions can vary according to raw material, meaning the same suites of characteristics might not be generalizable across raw material types (see especially Proffitt et al., 2021b).

Longitudinal studies (e.g., Pargeter et al., 2019, 2020; Stout et al., 2009) have followed the development of toolmaking skill and outcomes for novice learners during extended training periods. Individual skill acquisition also seems to involve optimization in terms of the application of functional actions and gestures for flake removals (Rein et al., 2014). These studies demonstrate that there is traceable improvement over time (e.g., decreases in core battering, decreases in production of angular fragments), but there can also be learning plateaus (Pargeter et al., 2019). At least for Late Acheulean biface production, full attainment of expert toolmaking skill is estimated to require considerable time investments (upwards of 200 hours, see Pargeter et al., 2019).

The earlier periods of skill development in living modern human novices are unusual, however, because they produce markers on the artifacts, which are not observed in the archaeological record (e.g., at Early Oldowan sites: cf. Braun et al., 2019; Pargeter et al., 2021; Stout & Semaw, 2006; Stout et al., 2009). The dissonance between human novice products and true archaeological artifacts is something of a conundrum. Some authors have proposed that early toolmaking hominins must have already possessed enhanced cognitive skills (e.g., related to know-how copying) in order to have achieved such refined skills (cf. Stout & Semaw, 2006; Stout et al., 2009, 2019). An alternative explanation pertains to the possibility that artifacts are not the outcomes of the solo actions of individual toolmakers, but

– just as are full assemblages – the products of averaged or summed actions of groups of individuals or time-separated individuals (consider also the arguments of Perreault, 2019). At least some experimental work has demonstrated that the combined actions of individuals with varied skill levels can essentially erase the evidence for novice workers (Ferguson, 2003; Sternke & Sørensen, 2007; see also Shipton & Clarkson, 2015 on biface resharpening effects on morphological types). Issues of raw material interactions, as well as the here described problem of individual versus 'group' agency, call into question the viability of direct comparisons between novice knapper products (especially in the standard experimental condition of 'single individual, single core') after only brief training periods with artifacts from excavated assemblages, which may be the time-averaged results of many individuals (see logic pro direct comparisons in e.g., Stout & Semaw, 2006; Stout et al., 2009, 2014, 2019; Toth et al., 2006).

Experimental archaeology might also improve our understanding of knapping's place in hominin life history. Various scholars have suggested that the production of knapping products began in childhood or adolescence (e.g., Ferguson, 2008; Hoegberg, 2008; Shea, 2006, Sternke and Sorensen, 2009). Relatively little attention has been devoted to finding the contributions of children to the Paleolithic record (Shea, 2006), with little to set a child novice apart from an adult novice. While replication of stone tools by adults – regardless of whether expert or novice – has a thorough precedent, there have been only anecdotal accounts (Ferguson, 2003; Sternke & Sørensen, 2007) and one replication study with child novices (Sternke & Sørensen, 2007). Human children have demonstrated at least limited capacities for producing stone tool types (though lacking in strength and sensorimotor coordination) and have also been observed to be capable of innovating bipolar techniques after exposure to finished stone tools and observation of freehand knapping.

A growing body of work has focused on contextualizing the specific mechanisms and processes of learning in prehistory, particularly with regard to the evolution of modern human capacities for learning and gestural and verbal language (Toth & Schick, 1994). Results and conclusions from these studies have been largely inconclusive and contradictory, with only limited differences between conditions in some cases, whereas in other studies, e.g., verbal teaching was more advantageous or gestural teaching was more efficient (see results and discussion in Cataldo et al., 2018; Lombao et al., 2017; T. Morgan et al., 2015; Putt et al., 2014; Ohnuma et al., 1997). At least part of the discrepancies between all of these studies might be the consequence of finer details in the experimental design (e.g., the kinds of demonstrations and teaching opportunities presented), the relatively short testing periods (reviewed in Pargeter et al., 2019), or the different technologies being replicated and raw materials being utilized (see again Proffitt et al., 2021b for an illustration of the interactions between novice-level experience and different raw material types). It is indeed possible that perceived differences in performance might level out in the long run, with learning processes being ultimately inconsequential for end-state mastery of knapping skill. Contra claims to the contrary (e.g., Lombao et al., 2017; Shipton, 2020; Toth & Schick, 1994), acquisition of toolmaking skill appears in all known testing conditions, including those perceived as 'minimally' necessary to toolmaking (often, emulation is treated as the absolute minimum; T. Morgan et al., 2015). Innovation of knapping techniques appears also in true baseline conditions (without any cultural transmission of know-how, including even the previously assumed minimum of emulation), thus showing – as a matter of proof-of-concept – that said cultural transmission of know-how is not necessary for modern humans and may also be unnecessary for extinct hominins (see Snyder et al., 2022; Tennie et al., 2016, 2017).

Neuroarchaeological approaches – the "application of neuroscience theory and methods to archaeological questions" (Stout & Hecht, 2015, p. 146) – have also provided intriguing insights into questions of cognition and learning in human and hominin prehistory. Neuroarchaeological studies (e.g., Hecht et al., 2015; Stout & Chaminade, 2007; Stout et al., 2011, 2015; Putt et al., 2017, 2019) have concentrated on concepts related to selective pressure on the evolving hominin brain (especially regarding co-evolution of language and technological capacities, or exaptation of technological systems to language). This line of research highlights the important of visuomotor functions and understanding of spatial relationships in the production of stone tools (e.g., Hecht et al., 2015; Putt et al., 2017; Stout & Chaminade, 2007; Stout et al., 2015). Neuroarchaeological studies have also demonstrated progressively intensifying cognitive demands in correspondence with evolving complexity of Lower Paleolithic technology (Hecht et al., 2015; Putt et al., 2017; Stout et al., 2015). Other relationships have not been replicated across studies. Earlier evidence for the coevolutionary relationship of technology and language (based on measurement of expert knappers; Stout et al., 2008) was not replicated by researchers studying human novices in verbal and nonverbal learning conditions (Putt et al., 2017). The latter favor the role of visual working memory systems in the evolution of technological complexity and human culture (Putt et al., 2017, 2019; Stout et al., 2015 also identified working memory as a major component to toolmaking). Finally, the cognitive demands of acquiring knapping skills have been shown to be so great as to result in structural remodelling of the brain regions currently identified as supporting human tool use (Hecht et al., 2015).

All experimental work shares a common observation, that of the importance – neigh, absolute necessity – of individual learning/practice to know-how acquisition (Harlacker, 2003; Nonaka et al., 2010; Pargeter et al., 2019, 2020; Parry et al., 2014; Rein et al., 2014; Snyder et al., 2022; Stout & Chaminade, 2007; Stout & Semaw, 2006; Stout et al., 2011, 2014). Again, the precise context – social or otherwise – of these learning opportunities is unclear (compare, e.g., Pargeter et al., 2019, 2021 with Snyder et al., 2022). Nonetheless, knapping skill is not something that can be gained from conceptual knowledge alone (Harlacker, 2003; Nonaka et al., 2010), but can rather be acquired even in the total absence of conceptual knowledge (Snyder et al., 2022). Naïve novices are not perfect masters of the toolmaking actions at their first attempts (Snyder et al., 2022; see also observations of Bargallo & Masquera, 2014; Bargallo et al., 2018; Stout & Semaw, 2006; Stout et al., 2009). Instead, they require *some* kinds of opportunities for toolmaking activities to occur and be repeated over time (whether that is a more ecological basis as in Snyder et al., 2022 or the concerted, deliberate practice assumed by, e.g., Stout & Hecht, 2017; Stout et al., 2011, 2014; Pargeter et al., 2019, 2020, 2021). It is only through these repeated events of hands-on practice that previously beginner toolmakers can become experts (Pargeter et al., 2019, 2020; Stout et al., 2009), even if that process seems largely unidentifiable in the archaeologically record (Ferguson, 2003; Stout & Semaw, 2006; Stout et al., 2009).

Epistemic problems in (cognitive) experimental archaeology

Problems of analogy

Analogic reasoning is a fundamental part of replication experiments in archaeology (Stout, 2005). The basic premise of this area of research is simply: artifacts produced in the present with processes in the present can essentially be compared with artifacts from the archaeological record to better understand processes in the past (e.g., Putt et al., 2017). We study human novices because we intend to make analogies with hominin toolmakers, whether that means developing inference points to estimate the skill level of those hominins (e.g., Stout & Semaw, 2006) or learning about the acquisition process itself and the cognitive mechanisms it supposedly necessitates (e.g., Stout et al., 2009) or with some other research question in mind. We build these analogies by necessity, since it is not possible to study the toolmaking behaviors of extinct hominin populations directly (cf. Harlacker, 2003; Stout, 2005) and analyzing artifacts alone is insufficient as a methodology for constructing middle range theories concerning the past (processes are not always obvious from the end-products, especially when equifinality might be at play; Perreault, 2019).

Whereas analogic reasoning of physical processes (like fracture mechanics) is relatively unproblematic given the reasonable assumption of consistency of underlying processes across time (*uniformitarianism*; cf. Lin et al., 2018), we cannot totally assume uniformitarianism between biological, cognitive, and culture processes in the past (e.g., social learning mechanisms, working memory, reduction sequences) and processes involved in the contemporary manufacture of stone tools. In the scope of the Mesolithic or Paleoindian assemblages (e.g., Ferguson, 2003; Sternke & Sørensen, 2007), our analogies might be *reasonably* trustworthy as the species producing them originally was also *H. sapiens* (though still consider the issue of culturally learned cognition; see Heyes, 2018). With increasing evolutionary and genetic distance from living humans, the reliability of these analogies degrades regardless of underlying processes. We cannot assume, irrespective of how convenient it would be, that early hominins had similar cognitive skills (and for similar underlying reasons) to our own (cf. Corbey et al., 2016; Hovers, 2012; Killin & Pain, 2022; Ohnuma et al., 1997; Putt et al., 2017; Stout et al., 2009; Tennie et al., 2016, 2017; Schick & Toth, 1993), especially seeing as these traits were likely under selective pressure for the last few million years and are thus represent potential confounding variables in their current derived states (see Hecht et al., 2015; Putt et al., 2014; Stout, 2006).

Our knowledge of brain function and cognition in the evolution of hominins is relatively constrained due to lack of truly direct evidence of said phenomena (e.g., Stout, 2006), but there are at least some identifiable differences – differences which might affect or alter behavioral and cognitive processes involved in both toolmaking and skill acquisition – between living humans and extinct counterparts that ought to be considered when conducting experimental knapping studies and interpreting the results thereof. Brains in living humans are larger absolutely and proportionally than those of apes and earlier hominin species (with the exception of Neandertals and some Middle Pleistocene hominins) and are also distinctive as far as the structural organization, characterized by a uniquely globuralized shape (e.g., Bruner et al., 2017; Hublin & Changeaux, 2022; Neubauer et al., 2018; Ponce de León et al., 2021). Indeed, the candidates for the earliest toolmakers (i.e., australopithecines) had brains with an ape-like brain organization (Gunz et al., 2020), with more modern brain organization appearing long after the first stone tools (Ponce de León et al., 2021). Similarly, post-cranial anatomy also differs across taxa and could have had a potential influence on the flexibility or constraints in behaviors, including indirectly via embodied cognition (Harlacker, 2003; Marzke & Shackley, 1987; Putt, 2015; Stout & Semaw, 2005; Stout et al., 2009; Williams et al., 2010; Williams-Hatala et al., 2018).

Additionally, life histories of early stone-knapping hominins were not identical to our own, and life history in turn may be connected to other factors like learning and social structure (cf. Sterelny, 2020). It is unclear what living models offer the best perspective on hominin development and learning, and the Paleolithic record offers no direct evidence in the way of when hominins would have started learning how to knap and how long this skill acquisition would have lasted. Adolescent apprenticeships in adze-making (Stout, 2002), child knapping experiments (Ferguson, 2003; Sternke & Sorensen, 2007), and the extensive acquisition (or perhaps mere maturation) period for nut-cracking in chimpanzees (Boesch and Boesch-Achermann, 2000; Inoue-Nakamura & Matsuzawa, 1997; Matsuzawa, 1994) represent just a few heavily contrasting, though equifinal, possibilities. Postnatal brain development, as one measurable component of life history and a learning-linked phenomenon, suggests that humans have a unique pattern of growth when compared with Neandertals (Gunz et al., 2010, 2012) and that *Australopithecus afarensis* – possibly the progenitor of the Dikika bone marks (McPherron et al., 2010) and the earliest stone tools (Harmand et al., 2015) – had a chimpanzee-like fast life history and human-like protracted brain growth (Gunz et al., 2020; Hublin et al., 2015). Should brain development and adult brain anatomy be acceptable proxies for other life history and cognitive factors, then perhaps *ape-based* learning models would be more appropriate than modern humans for 'childhood' learning and transmission in the first toolmaking hominins (though, again, apes would be imperfect models; cf. Bandini et al., 2022). Finally, the social contexts that we see in human stone toolmaking and craft learning, like apprenticeships (e.g., Bamforth & Finlay, 2008; Stout, 2002) and hunter-gatherer subsistence techniques (e.g., Shea, 2006), cannot be treated as representative for the Paleolithic context given the sheer lack of tangible evidence for their existence back then (see Putt et al., 2017; Tennie et al., 2017).

Most studies involve participants from WEIRD backgrounds (and in small sample sizes; see discussion in Gabrić et al., 2021; Killin & Pain, 2022; Pargeter et al., 2019, 2021), which are not only unlikely to be representative of cognition for all living human populations (Henrich et al., 2010) but also unlikely to be representative of the lifeways, behavioral patterns, and also learned cognition that could be assumed as typical for extinct hominins and earlier humans for the vast majority of prehistory (Bandini et al., 2022; Killin & Pain, 2022). Humans, especially WEIRD humans, are extremely reliant on the use of specific cognitive mechanisms like imitation and language, which they use to adopt and transmit know-how (see Bandini et al., 2022) and which are seemingly lacking from other primates and perhaps also hominins (Dean et al., 2013; Tennie et al., 2009; Tomasello, 1999). Indeed, learning environments during development and learning of specific behaviors can shape even brain activation and plastically restructure the way our minds function in relation to certain behaviors and technologies (Hecht et al., 2015; Heyes, 2019; Killin & Pain, 2022; Putt et al., 2017, 2019). On top of that, WEIRD-typical learning preferences can also affect the abilities of both learners and teachers to engage in certain conditions (see Pargeter et al., 2021; Putt et al., 2014; Ohnuma et al., 1997). Past learned information has the capacity to upregulate or downregulate abilities shown in cognitive knapping experiments (cultural norms in WEIRD societies against breaking objects might impact knapping outcomes versus what one might see in, e.g., Ethiopian pastoralists that regularly engage in more 'haphazard' stone fracturing activities; see Snyder et al., 2022). A modern human mind is therefore a very special thing that cannot be readily applied (i.e., without acknowledged caveats) to past contexts, and thus, just because we have hominin artifacts, does not mean we can automatically say that there was 'human' cognition (Tennie et al., 2017).

Competing approaches to evolutionary cognition

Traditionally in cognitive archaeology, researchers have often followed the logic of the minimal necessary competence. In essence, the focus in this approach is on the cognitive capacities that are minimally necessary to achieve the production of an artifact (e.g., Wynn, 1979, 1988; Wynn & Coolidge, 2004). This has been applied, for example, to determine "what essential preconditions (minimal capacities) does [the Oldowan's] initial appearance imply" (Stout, 2006, p. 287). Though more cognitive capacities cannot be completely excluded and the "full extent of the makers' mental capabilities" are unlikely to be reflected by any single artifact (Stout & Semaw, 2006), the intent here is to not assume anything more was going on than can be determined by models from the cognitive sciences or from empirical testing. This approach has received criticism by some researchers (Killin & Pain, 2021, 2022), based on the presumption that toolmakers "may well have exercised their most sophisticated thinking sophisticated thinking in ways that are archaeologically invisible"

(Killin & Pain, 2021, p. 3). Killin & Pain (2021) further critique Wynn & McGrew's (1989) interpretation of Oldowan tools being indicative of ape-like cognition and their call for evidence- rather than assumption-based conclusions in cognitive archaeology, posing the example of a Bronze Age village that lacks direct evidence of the use of symbols being interpreted as non-symbolic despite wider use of symbolism in the Bronze Age. Here, it is tempting to say that the same metaphor could apply to other periods (e.g., the Oldowan and Acheulean), but in cases of evolutionary cognition this is simply not a functional argument.

In terms of direct behavioral remains for much of the Paleolithic, all we have for study is stone tools, and when there is extensive temporal and geographic conservation in stone tools and their reduction sequences, it becomes rather incongruent to assume that any more cognitive complexity was taking place than what the stone tools and replication experiments thereof can tell us. As living modern humans are the products of generations and millenia of evolution *post facto*, ideal conditions for learning and optimized cognitive capacities of living modern humans are therefore a poor frame of reference. Humans are not the same as extinct hominins and as such, we should indeed focus on the minimal conditions and cognitive capacities for achieving toolmaking and tool use tasks. As a proof of principle (Snyder et al., 2022), it can be assumed that "extinct hominins producing the same tool types and using the same operational sequences as modern humans likely possessed at least the minimum cognitive operations that modern humans use to complete the task" (Putt et al., 2017). Although Killin & Pain (2021) are correct in their suggestion of using multiple approaches to resolving issues of minimal necessary competences, we reject their dismissal of parsimony for ascertaining the 'best guess' of several possible answers to determination of cognitive faculties of non-humans and ancient hominins.

In some cases – and contrary to the logic of parsimony – researchers still assume beyond the minimal competences required for a specific artifact or technology, very often the case in the sphere of social learning where, e.g., emulation (results copying) has been shown in various cases to transmit know-how (also a priori assumed to be the minimal possible condition for know-how development) and yet more evolutionarily 'advanced' copying social learning mechanisms are implicated instead in Oldowan and other stone tool industries (cf. T. Morgan et al., 2015; Nielsen & Whiten comment on Tennie et al., 2017). Too often, some cognitive traits are simply taken as a given and certain social learning mechanisms are assumed to always have been present (or simply, because humans do it, it suffices; e.g., Pargeter et al., 2019), ignoring the possibility that these were either not present (not even testing for possibility these were unnecessary; see Snyder et al., 2022; Tennie et al., 2016, 2017) or did not function in some species-specific/evolutionarily-distinct way (see Pargeter et al., 2021). This is a tautological fallacy, as, if a mechanism like copying or teaching is already implicit to the approach, the results and interpretation of the test are almost bound to support the mechanism's value for the existence of behavioral know-how (Tennie et al., 2017). Often, this is based on interpretations that cultural transmission of know-how and other mechanisms are present in non-human apes (e.g., White et al., 1999), but there are grounds to doubt the cultural transmission abilities of these species (reviewed in Bandini et al., 2020; Tennie et al., 2020). This means we cannot simply take copying (e.g., a minimal level of emulation or imitation; cf. T. Morgan et al., 2015) as the axiom for hominin behavior and toolmaking. In other cases, traits – such as particular grips, recruitment of musculature, affected brain pathways, etc. – that are identified as being used for toolmaking tasks by modern humans are suggested to be target areas for selection during the evolution of the hominin lineage rather than necessarily requiring human-like versions of those traits in earlier species (e.g., Hecht et al., 2015; Marzke et al., 1998; Stout, 2005; Williams-Hatala et al., 2018).

In increasing frequency, researchers have made efficiency-based argumentations for the presence, operationality, and necessity of specific cognitive mechanisms in prehistory. Most of this argumentation is based on comparison between markers of learning exhibited by human novices in varied conditions of learning and communication. The standard logic is that more efficient or effective communication variants, social learning mechanisms, and general learning contexts are candidates for being under selective pressure (as mentioned above) or being specifically necessary for the transmission of technical know-how (cf. Cataldo et al., 2018; Lombao et al., 2017; T. Morgan et al., 2015; Putt et al., 2014; Shipton, 2020). Efficiency arguments, however, are deeply issue-laden, not only for directly defying the logic of minimal competences.

One major issue is that using 'efficiency' as the basis for arguments about past hominin populations relies on inferences for strong similarities in – nigh identical – cognition of hominins and humans. However – also acknowledging the differences in cognition between apes and humans and varied comparisons between apes, humans, and hominins (Bandini et al., 2020, 2022; Dean et al., 2012, 2013; Stout et al., 2009; Tennie et al., 2009, 2016, 2017, 2020; Toth et al., 2006; Wynn & McGrew, 1989; Wynn et al., 2011) – it cannot be assumed that what is cognitively or behaviorally efficient for humans is also efficient for hominins. For instance, if proto-language evolved as a non-technological adaptation (e.g., related to social coordination or some other group-level activities; see recent evidence regarding communication in apes; Mine et al., 2022), then proto-language may not have been as efficient for transmitting technological know-how as gestural and verbal language have been measured and interpreted to be in living modern humans learning to make stone tools (compare with conclusions of Cataldo et al., 2018; T. Morgan et al., 2015; Lombao et al., 2017; Putt et al., 2014; Ohnuma et al., 1997). Similar arguments might be made for other cognitive mechanisms, such as imitation and teaching. Efficiency is also emphasized, because toolmaking – though we do not know just how significant of a role it played in the lives of our predecessors (see Shea, 2017; Stout, 2006) – is thought to have been a practice upon which the success and survival of hominin individuals and groups would have depended (Pargeter & Shea, 2019; Pargeter et al., 2019, 2020; Shea, 2017, 2019; Toth & Schick, 2018). Although this may well be true, the centrality, neigh unavoidable necessity, of stone toolmaking could still be overestimated, due to erroneous applications of human life history and apprenticeship models (see above) and the lack of preservation of non-stone tools that would have contributed to survival as well (cf. Ambrose, 2001; Bandini et al., 2022; Gabrić et al., 2021; Haslam et al., 2009; Hovers, 2012; Rolian & Carvalho, 2017; Toth & Schick, 2009; see also Dusseldorp & Lombard, 2021).

To expand on these issues with using arguments by efficiency, it might be useful to consider a few more examples. First, if one were to test toolmaking skill acquisition in deaf and hearing-abled individuals (as proposed by Toth & Schick, 1994), it would be rather silly to assume that no toolmaking would occur in deaf individuals (thereby placing question marks on claims for verbal language to have been absolutely necessary in early stone toolmaking industries). Additionally, one might predict that mechanisms for gestural language (especially, due to enculturation of deaf persons in said medium) and for imitation might be even more efficient than in hearing-abled individuals; the point here is that improved efficiency of a mechanism in any one population or study group cannot simply be assumed to directly apply to hominins as well. One might also consider the following thought experiment: we could prepare an experimental apparatus consisting of two testing groups, engineers and non-engineers. If the engineers are significantly more efficient or effective at learning to make stone tools and one applies the logic from these studies, then the conclusion would follow that the target hominins had also studied engineering – and even had to study engineering. The logic of the engineering example is faulty, and so is the same logic applied for various social learning mechanisms and language. As non-engineers can also learn to make stone tools (see also Ferguson, 2003), the higher efficiency of learning with an engineering degree cannot be used at all as an argument for necessity in prehistory.

In cases where multiple (cognitive) processes can produce the same artifact outcomes (and having only those artifact outcomes in the archaeological record), parsimony favors the least 'complex' (i.e., with the fewest inferred evolutionary steps) of the equifinal possibilities. Considering that hominins would have potentially had their entire lives to individually attain skill rather than mere minutes or hours as in these experiments (cf. Eren et al., 2016), observed short-term advantages in terms of efficiency alone could actually be fully inconsequential for hominins who had years to learn the techniques by supposedly less efficient means. Here, the travelling friend metaphor of Perreault (2019, p. 38) is particularly illustrative: a friend from California messaging to say they will visit you in New York arrives either the next day or the next week. In the former scenario, there is one suitable explanatory process (i.e., mode of transportation, by plane), while in the latter, several different processes could be true (e.g., by plane, train, car, helicopter, or combination thereof). This is the same for transmission experiments where multiple social learning mechanisms are effective at reproducing the elements from the archaeological record despite the variation in presumed and/or measured efficiency and related time-costs.

Given the serious epistemological problems involved in efficiency-based approaches to evolutionary cognition, we advise for a return to more strict minimal necessary competences approaches for evaluating the cognitive implications of prehistoric stone tools. Our primary methodology in this regard is the use of so-called island tests (i.e., true baseline experiments for targeted know-how capabilities; cf. Bandini et al., 2020; Tennie et al., 2016, 2017, 2020; Tomasello, 1999). In these true baselines, experimenters control for all opportunities for the direct cultural transmission of know-how, in essence: there are no possibilities for mechanisms like imitation, language, and teaching of know-how. Given differences between humans and hominins, this is all about proof-of-principle; if humans can do it without copying social learning for instance than it is conceivable that a similar case is true of hominins, especially given similar patterns in other species (Motes-Rodrigo et al., 2022; Westergaard & Suomi, 1994a, 1995a).

Natural quasi-baselines have occurred in the form of early replicative knappers (e.g., Callahan 1979). Archaeologists and hobbyists in the 20th century re-innovated stone tool technologies almost completely independently and without observing how *H. erectus* or other toolmaking hominins did so. At least two anecdotal accounts show that children can re-innovate anvil-based knapping without observation of that specific technique (Ferguson 2003, Sternke and Sorensen 2007). In our most recent study (Snyder et al., 2022), we invited participants to a problem-solving study and ask them to use provided raw materials (i.e., valid material for conchoidal fracture and flaking) and nothing more to access monetary rewards in a puzzle box.

The unavoidable drawback of this approach is the potential enlistment of completely or partially non-naïve individuals into the test cohort. To overcome this, participants were evaluated afterward via a naivety questionnaire. This questionnaire checked for a variety of information sources to which individuals might have previously been exposed (see Table 4). A truly exhaustible list of sources is nigh impossible, so participants were also given the opportunity to write in other applicable experiences in their memory. Not only does this questionnaire serve to filter out non-naïve individuals from further testing, but it also allows for the incorporation of naivety levels and pre-study information types into the analysis of the wider experimental dataset, meaning the influence of these factors on the overall learning model across individuals can be evaluated. We believe that this approach is the most effective

Table 4 Naivety ranking system developed in conjunction with a post-study questionnaire on previous experiences. This serves as a practical tool for experience level-based data treatment, as well as a basis for exclusion from subsequent testing and analyses. In our experiments, level 5 would be totally excluded.

Level	Description	Data treatment
0	Participant is *totally* naive In this case, answers "No" to all pertinent questions in questionnaire	Ideal subjects: individual cases described in detail and included for final analysis
1	Participant has conceptual knowledge, i.e., is aware of the existence of sharp stone tools *E.g.,* Stone tools mentioned in a primary school history class, documentary	Not perfectly naïve; individual cases described in detail and included for final analysis
2	Participant has seen a stone tool or a depiction of a stone tool *E.g.,* Museum visit, textbook figure	
3	Participant has seen the production of a stone tool *E.g.,* YouTube videos, live demonstrations	Included in the analysis, primarily to determine effects of non-naivety on task performance and end-products
4	Participant has direct hands-on experience with knapping *E.g.,* Knapping courses, other knapping experiments	
5*	Participant has theoretical knowledge of stone tools and their production *E.g.,* Participant is an archaeology student or career archaeologist	

thus far conceived for testing the knapping abilities of naïve novices, and our particular programme is the only of its kind to attempt to ensure the total naivety of the involved participants. In the end, we were able to validate many participants in our study (n = 22) as being technique-naïve, and two even turned out to be totally naïve (they had never before heard of stone tools as a concept). Despite this naivety to the technical know-how of toolmaking, these participants were capable of reproducing valid early knapping techniques and making artifacts resembling those from Oldowan assemblages. On this basis, we can exclude any mechanisms for direct cultural transmission of know-how from the minimal necessary competences for, e.g., Oldowan or Mode 1 toolmaking.

Further assumptions and fallacies in archaeological experiments

Expert bias can have a serious impact on all steps of the scientific process (Bisson, 2001; Eren et al., 2016; Marzke & Schakley, 1987; Shea, 2022; comment by Davidson on Tennie et al., 2017). Archaeologists often assume, for example, that the intentions and actions applied during actualistic toolmaking studies are the same intentions and actions that earlier hominins would have used for their own toolmaking activities (e.g., Schick & Toth, 1993). Most often, these inferences are based on earlier research that determined these to be viable for the replication of prehistoric tool forms and then repeatedly re-applied as the 'correct' methods in successive studies. In many cases, this may be – in fact, most likely is – fallacious (cf. Eren et al., 2016; Shea, 2022), as settling on one 'correct' possibility ignores the potential for equifinality of processes (Perreault, 2019). There are numerous processes at differing levels of explanation where this could be the case, including gestures, knapping techniques, hammer types, learning contexts and curves, cognitive mechanisms, and knapper intentions.

For example, some authors (e.g., Lombao et al., 2017; Schick & Toth, 1993) have claimed that totally naïve individuals are unable to apply the 'correct gestures' for achieving conchoidal fracture for flaking. Yet other studies show that inexperienced individuals are immediately capable of performing flaking tasks (e.g., Bargallo & Masquera, 2014; Bargallo et al., 2018). Learners have separately been shown to adopt intra- and interindividually varying gestures for (successful) knapping actions, in many cases even using gestures distinct from the ones demonstrated to them (Rein et al., 2014). Similarly, human novices have proven capable of innovating different techniques for achieving conchoidal fracture (Ferguson, 2003; Pargeter et al., 2021; Snyder et al., 2022; Sternke & Sørensen, 2007). Knapping techniques have been shown in some cases to be difficult to distinguish based on the quantifiable traits of products from, e.g., bipolar and freehand knapping (Byrne et al., 2016; Duke & Pargeter, 2015; Pargeter & Eren, 2017). Problems with identifying technique from the archaeological evidence can also be exacerbated if cores are modified by and reduction sequences involve more than one technique (e.g., Hiscock, 2015; B. Morgan et al., 2015; Snyder et al., 2022). Excessive faithfulness to the gestures and techniques used by expert knappers may then lead to an underestimation of the true diversity of methods employed in the past, as well as the estimated proportions that are derived from archaeological assemblages (see Bril et al., 2010).

Similarly, experimentally archaeologists (by necessity, in order to conduct replication experiments at all; see Pargeter et al., 2019; Stout et al., 2014) make assumptions about the intentions and goals of prehistoric toolmakers. In truth, we often do not know the precise goals of ancient toolmakers (see Gabrić et al., 2021; Stout et al., 2019). For example, the intentional production of an Oldowan chopper in a modern replication experiment is not representative of the cognitive underpinnings of tool manufacture if the interpretation of Oldowan tools as unintentional byproducts of least effort flaking holds true (see Gabrić et al., 2021; Moore & Perston, 2016; Snyder et al., 2022; Toth, 1985). Likewise, frequent debates center around the intentionality of handaxe morphologies, which some experimentation has shown to arise – at least in a 'crude' form' – from stochastic reduction sequences (Moore & Perston, 2016), but which are interpreted to be intentionally shaped objects in yet other experiments on Late Acheulean technology (Pargeter et al., 2019, 2020; Stout et al., 2014). By following methodologies that carry this inference of intentional shaping, other possible explanations for handaxe shapes, e.g., related to functionality (cf. Iovita & McPherron, 2011), are left unexplored and thus their alternative influences on factors of learning and cognition are also unstudied. Distinct intentions can potentially create the same stone artifacts (see additionally, the 'finished artifact fallacy'; Davidson, 2002), while also interacting variably with other cognitive and physical processes key to toolmaking; these possibilities require acknowledgment, first and foremost, as well as further experiments so that untested assumptions are not left untested.

The next issue regards 'natural' conditions in knapping experiments (discussed in Pargeter et al., 2019). Knapping experiments have been heavily criticized for the artificial nature of their methodologies and supposed disconnect from the reality of knapping in prehistory, including the relatively short testing times and the occasional use of materials that would have been inaccessible to extinct hominins (Pargeter et al., 2019, 2021). Yet, even those experiments utilizing so-called ‘natural’ learning settings, also suffer from issues with external validity. Key to this argumentation is the assumption that stone artifacts have a ‘tool’ status and would therefore have served a purpose beyond just their manufacture. However, in most knapping experiments, there is a total separation of toolmaking from the subsequent tool use (only in rare examples were participants asked to test their tools’ utility on supplied raw materials; Bisson, 2001; Stout & Semaw, 2006; Stout et al., 2009). Instead, knapping experiments either involve the removal of a prescribed number of flakes from a core (e.g., Nonaka et al., 2010; Oldowan condition in Stout et al., 2015) or the exhaustion of one or more cores in single bouts. This would not be a representative testing model of technologies wherein mobile toolkits were regularly used over many episodes (even over many years of activity), involving a fluid chain of episodes of toolmaking and tool use (i.e., make a flake, use a flake or core-tool, repeat; see Snyder et al., 2022). Viewing toolmaking as deliberate practice that occurred separately from use, perhaps even until ‘expert’-level performance was achieved, therefore exaggerates the degree of costs involved in learning to knap, when actually knapping as a regular activity integrated into ecological tool use (e.g., for extractive foraging) would have led to learning and skill improvement by proxy rather than with intentional investment (contra e.g., Pargeter et al., 2019, 2020).

Furthermore, archaeological sites are often accumulations and amalgams of long-term activities by multiple individuals, even potentially from unrelated groups or separate species (Perreault, 2019; Semaw et al., 1997; Stout & Semaw, 2006; contra Stout et al., 2019). By this notion, the actions of novices may have been erased or lost due to averaging effects of multiple agents involved in the same assemblages (e.g., Ferguson, 2003). This could explain the stark differences between archaeological materials and artifacts produced by human novices, whereby archaeological assemblages seem to largely lack a signal that one might associate with the exclusive work of beginners or early learners (e.g., Ferguson, 2003; Pargeter et al., 2021; Snyder et al., 2022; Stout & Semaw, 2006; Stout et al., 2009). As a general rule, this shows us that direct variable-by-variable comparisons between experimental and archaeological data can be uninformative unless immediately contextualized by the conditions under which the experimental data was collected and the research questions at hand.

Prospects and conclusions

The study of novices in replication experiments is an increasingly important aspect of research into hominin cognitive evolution. Novices, by the nature of their inexperience, are the optimal subjects for tests on skill acquisition, learning, cultural transmission, and innovation. More specifically, researchers in this field have attempted to acquire insights into how hominins acquired the skill of stone toolmaking, whether this is understanding the skill itself or the evolutionary processes that are tied to toolmaking (i.e., the evolution of language, social learning and hominin brains, the emergence of the earliest knapping techniques). These insights require the drawing of analogies between modern novice knappers and prehistoric equivalents; these analogies are burdened by clear and obvious faults but are nonetheless essential in some shape or form.

Despite all of issues with inference-making and analogies between modern and prehistoric humans, our recommendation is not that experiments with modern knappers should be abandoned. Instead, a more nuanced approach is required, as experiments still can produce tangible insights into the evolutionary past – even if these are 'only' educated guesses and middle-range theories and not necessarily indisputable fact (Geribàs et al., 2010; Ohnuma et al., 1997; Putt, 2015; Schick & Toth, 1994; Stout, 2005; Stout & Chaminade, 2007; Stout & Semaw, 2006; Stout et al., 2015; Tennie et al., 2017; Schick & Toth, 1993; Williams et al., 2010; Williams-Hatala et al., 2018, 2020). Well-constructed experiments are certainly better than none (just looking at modified rocks is not sufficient to inform us about the processes that led to their existence; Binford, 1972; Killin & Pain, 2022), and living humans (and non-human primates) are still the closest phylogenetic neighbors of extinct hominin species, which would suggest that they are serviceable as the best models for extinct hominins in experimental studies (Bandini et al., 2022; Snyder et al., 2022).

Good scientific practice in experimental archaeology should mean clarity about any (especially untested) assumptions involved and the justifications for making such assumptions. Often, as a part of designing and conducting controlled experiments, assumptions have to be made in the testing model, simply as a way of maintaining variable control (Eren et al., 2016; Lin et al., 2018; Pargeter et al., 2019, 2020). Just as well, experimental design also includes a negotiation or balancing act between variable control and external validity (Eren et al., 2016; Line et al., 2018; Pargeter et al., 2019; Snyder et al., 2022).

We should examine our priors as researchers (Bandini et al., 2022; Eren et al., 2016) and consider models that can provide alternative insights that are equally valid for our understanding of what happened in the past. This means not limiting consideration to the 'correct answers' previously defined by expert archaeologists and knappers (Bisson, 2001; Eren et al., 2016; Marzke & Schakley, 1987; Shea, 2022; comment by Davidson on Tennie et al., 2017) but also incorporating data for a wider range of possibilities for gestures, techniques, and other behavioral and cognitive traits and processes. Such insight can be gained from ethnographic studies of living knapping groups (cf. Binford & O'Connell, 1984; Marzke & Shackley, 1987; Toth, 1982) or from experimental study groups (see Rein et al., 2014; Williams-Hatala et al., 2018). In this regard, studying totally naïve individuals in island tests can be an especially fruitful exercise, as these individuals can display behaviors that are absent from the archaeologically canon and yet are valid ways to make and use cutting tools (Snyder et al., 2022). Island tests (e.g., for determining the necessity for cultural transmission of know-how; Snyder et al., 2022; Tennie et al., 2016, 2017; Tomasello, 1999) may not be needed for absolutely every research question, but the (sometimes unexpected behavioral) data produced therefrom can be useful for formulating new testable hypotheses for follow-up, non-island-test experimentation. Generally, we would also recommend more thoroughness in determining the past experiences of participants as any of these could act as confounds in the results and participant bias can be just as much of an issue as expert/experimenter bias (e.g., Stahl, 2008; naivety differences did not lead to differences in the expression of behavioral know-how or the resultant artifacts, however, in Snyder et al., 2022).

Overall, given the persistent issues of equifinality, we should adopt an anarchistic attitude, testing all conceivable possibilities, even when 'obvious' answers have already been identified. Just as well, this involves the accumulation and triangulation of data from varied experimental approaches, study populations, and research objectives (e.g., Harlacker, 2003; Killin & Pain, 2021, 2022; Tennie et al., 2017). Only after multiple possibilities have been tested (perhaps not exhaustively, but certainly more intensively than is presently standard practice) is it then possible to start comparing and looking for best 'causes' (e.g., by parsimony; see Snyder et al., 2022; Stout et al., 2019; Tennie et al., 2017). Data from the studies reviewed in this manuscript should be complemented by further experimentation, as well as studies from other domains.

Ethnographic studies, another popular source of inferences about prehistoric toolmakers (e.g., Bamforth & Finlay, 2008; Duke & Pargeter, 2015; Ferguson, 2003; Shelley, 1990; Sternke &

Sorensen, 2007; Stout, 2002) and field experiments have their own limitations (i.e., related to experimental control; Stout, 2002), while controlled novice knapping studies thus far have been overwhelming WEIRD-focused and therefore lacking in generalizability (see Bandini et al., 2022; Henrich et al., 2010; Killin & Pain, 2022; Wynn, 1985). Future work should thus focus on controlled experimentation with further non-WEIRD populations, including (especially) groups that do not practice knapping or which have ceased doing so in recent generations (Killin & Pain, 2022; Snyder et al., 2022; Wynn, 1985). Besides non-WEIRD populations, additional research should be conducted on children, in order to expand on the relatively meager amount of data that has been procured thus far (cf. Sternke & Sørensen, 2007). Research with children would be beneficial for studying ontogenetic influences on skill acquisition and development, as well as in achieving greater control over indirect transmission of know-how than is possible when testing adults (see potential concerns outlined in Bandini et al., 2022). Further insights from non-human primates are also important for triangulation and contextualization of results from novice human knappers, with humans and primates (especially apes) being differentially suitable models for extinct hominin toolmakers (see arguments above; see also Bandini et al., 2021c, 2022; Davidson & Noble, 1993; Putt, 2015; Putt et al., 2014; Schick et al., 1999; Stout & Chaminade, 2007; Stout et al., 2009; Toth et al., 2006). Besides the aforementioned avenues of research, modelling (e.g., agent-based modelling) can be used to study processes at (larger) scales that are not so readily studied with experiments.

Finally, we need more longitudinal data under different conditions beyond those that benefit from "efficient" learning of typical WEIRD individuals (compare with Pargeter et al., 2019, 2020). Most of the present evidence from novice knapping studies is based on short, restricted training periods that may only represent early learning phases that are potentially untraceable in the archaeological record (cf. Pargeter et al., 2019; Stout & Semaw, 2006; Stout et al., 2009). Because the archaeological record only presents us with process *outcomes*, it is necessary to conduct longitudinal studies under differing conditions (e.g., learning conditions) until the interpreted 'expert'-level performance known from early stone tool sites can be theoretically reached (if it can be reached at all under any and all conditions; see Pargeter et al., 2021; Snyder et al., 2022; Stout & Semaw, 2006; Stout et al., 2009).

Naturally, not all cognitive and/or experimental archaeologists may not agree with our conclusions. Nonetheless, these should serve as food-for-thought, helping reflect on what it means to use living modern humans as a model organism in the place of pre-modern hominins and how using models affects how we conduct research and what interpretations we can really derive from the results of said research.

Acknowledgments

The project STONECULT was funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 714658). Special thanks go to Jonathan S. Reeves for his valuable input during the manuscript drafting process. Thanks to Patrick Cuthbertson for making the map of novice knapper experiments.

III. An overview of standardizable raw materials for controlled human knapper experiments

Explanation of candidate contribution to collaborative work

This chapter is based on collaborative work. The text represents the current version of the following manuscript:

Snyder, W.D., Boysen, D., Orellana-Figueroa, J.D., Tennie, C., & Reeves, J.S. (in prep). An overview of standardizable raw materials for controlled human knapper experiments.

Conceptualization: W.D.S., D.B., J.S.R. Investigation: D.B., W.D.S., J.S.R., C.T., J.D.O-F. Visualization: J.D.O-F., W.D.S., D.B. Formal analysis: W.D.S., D.B. Writing (original draft): W.D.S., D.B., J.D.O-F. Writing (review and editing): W.D.S., D.B., J.S.R., C.T., J.D.O-F.

Experimental replication of stone tools is an important method for understanding the context and production of prehistoric technologies. In scientific undertakings, experimental control is a valuable and necessary means of ensuring that confounding variables are not influencing the outcomes of the study. One way that researchers can exert control in knapping experiments is to standardize the knapping blanks (in size and/or form) and raw materials that are provided to the knappers. Though standardization of materials is already part of archaeological praxis, specific protocols, let alone comparisons between standardization techniques, are rarely openly reported. We investigated a variety of techniques and raw materials (e.g., hand-knapped flint, machine-cut basalt, manufactured glass, and porcelain) and evaluated them for their validity and reliability. Here, we have outlined the raw material tests we performed, providing information on the individual approaches, as well as comparisons between the techniques and materials according to both validity and reliability, along with their costs and our own recommendations. This text is intended as a serviceable guide on raw material standardization for knapping experiments, including previously pursued avenues and means that are as-of-yet undescribed in the experimental archaeology literature. We include additional considerations for techniques that we were unable to test, or which are presently not widely available and/or affordable. Future potential in this field would benefit from advances in the relevant technologies and in methodological approaches.

Introduction

Archaeologists have a long tradition of replicating the processes and end products associated with stone toolmaking (Johnson, 1978). Experimental replication of stone tools has generally been performed in order to better understand how those tools would have likely been produced in the prehistoric past (e.g., Johnson, 1978; ; Schick & Toth, 1993; Toth, 1985; Toth & Schick, 1994). Such experimental replication of prehistoric artefacts has been used to elucidate processes like the mechanics of conchoidal fracture (e.g., Dibble & Whittaker, 1981; Dibble & Režek, 2009; Dogandžić et al., 2020; Li et al., 2022), the reduction sequences of specific artefact types (e.g., Toth, 1985; Moore & Perston, 2016), and the cognitive mechanisms that guide the production of stone tools and the acquisition of the toolmaking skill (e.g., Putt et al., 2014; T. Morgan et al., 2015; Lombao et al., 2017; Pargeter et al., 2019; Snyder et al., 2022). Though early replication attempts were non-empirical (i.e., falling more within the scope of the humanities), modern experimental archaeology has placed a stronger emphasis on controlled experiments for the study of the physical as well as cultural and biological processes inherent to knapping behavior (cf., Whittaker, 1987; Eren et al., 2016; Lin et al., 2018).

Stringent experimental control can help to isolate variables of interest from potential confounds. If the independent variables of interest are related to cognition (i.e., skill level of the knapper or learning mechanisms), for example, extraneous physical variables should be controlled for by the experimental protocol (e.g., by constraining raw material quality, core shape, core size; cf. Lombao et al. 2017). Standardized blanks can also be used outside the experimental setting, especially as teaching tools in the context of public outreach or the training of new archaeologists (cf. Shea 2015). One particularly useful means of experimental control is the standardization of raw material types and blank forms (including blank size). Researchers have previously created and utilize standardized blanks (in some instances, preforms) in replication experiments (e.g., Sheets & Muto 1972, Dibble & Whittaker 1981, Dibble & Pelicin 1995, Khreisheh et al. 2013, Režek et al. 2016, Speer 2018, Dogandžić et al. 2020). Standardized blanks can be made from naturally occurring or synthetic materials - so long as they can fracture conchoidally - and are somehow shaped to follow a specific design concept. Standardizing blanks minimizes the effects of raw material geometry on any outcomes (artefactual, behavioral, or otherwise). Standardization does not always occur in nature (depending on geological formation and erosive processes). As such, standardized blanks have been argued to be less ecologically valid, with confounds of blank variability being overcome by using considerably large samples of blanks (cf. Pargeter et al., 2021). Nonetheless – though blank standardization might reduce outcome variability in some sense – it is still useful for pursuing certain research questions and, ultimately, is still reasonably valid so long as the standardized form is similar to or falls within the range of theoretically possible blank forms (e.g., Toth, 1985; Snyder et al., 2022).

Essentially, any attempt at standardizing knappable blanks should be focused on generating blanks that are consistently of the same size, shape and material (except in those cases where size, shape and material of the blanks and resultant cores are variables of interest, in which case variation of said variables would still follow the same protocols that we describe below). Here, we define two main components that must be kept in consideration when selecting raw materials and refining standardization techniques for a replication study: *reliability* and *validity* (Lin et al., 2018). Reliability refers to the capacity for a raw material to be standardized - otherwise, to be repeatedly created/shaped to a specific shape and size - and, by extension, the relative ease and efficiency with which the raw material can be standardized. Validity refers to the appropriateness of a raw material and/or blank form as an analogue to artifacts and materials from the archaeological record (i.e., external validity *sensu stricto*, e.g., Lycett & Eren, 2013; Eren et al., 2016). The most basic test of external validity is whether a raw material and/or blank form can be flaked via conchoidal fracture (cf. Cotterell & Kamminga, 1987; Dogandžić et al., 2020). In some cases, more precise external validity is required, meaning standardized blanks must be suitable according to the particular technologies, time periods, or sites being investigated (e.g., a study on the cognitive abilities of a particular population would best be carried out with the materials available to that population, as in Stout & Semaw, 2006). Although reliability and validity are not necessarily mutually exclusive, it is still possible that highly reliable techniques and materials are not particularly valid, and *vice versa* (see Discussion). With further practical and economic restraints on which materials and techniques would be useful or accessible, the actual application of raw material standardization to experimental designs is based on a negotiation between a variety of different influencing factors.

Standardization processes can be either *reductive* or *additive* (Ferguson, 2003; Schillinger et al., 2014). This useful distinction refers to how (via manual actions or machinery) the blanks are brought into shape from the original raw material form (e.g., the transformation of porcelain powder into slip followed by pressing into a mold is additive while the grinding down and sawing of stone is reductive). Additive techniques are generally more reliable than reductive ones as they more often produce blanks that either exactly match or are close to the ideal form. Most require the mixing of ingredients (such as those for concrete or powdered clay and water for porcelain) and a mold to encapsulate the developing blank. Additive techniques, though reliable and often capable of producing raw materials that conchoidally fracture, do not - strictly - result in raw materials that would have been available ecologically, therefore representing a potential loss in absolute validity. Steps can still be taken during production to make these artificial materials resemble the naturally occurring ones as close as is feasible (e.g., varying chemical formulas of ceramic powders, heat curves, generating heterogeneity in porcelain or concrete). The most basic reductive approach for making semi-standardized blanks would be knapping by hand to shape the material until it more-or-less fits the prescribed size and shape parameters (as in, e.g., Bisson, 2001; Bril et al., 2010; Bril et al., 2015; Rein et al., 2013; Nonaka et al., 2010; Bandini et al., 2021c; Motes-Rodrigo et al., 2022). More advanced techniques that achieve greater reliability involve the use of machinery such as diamond grinders, saws (in our case, also operated by hand), and rock

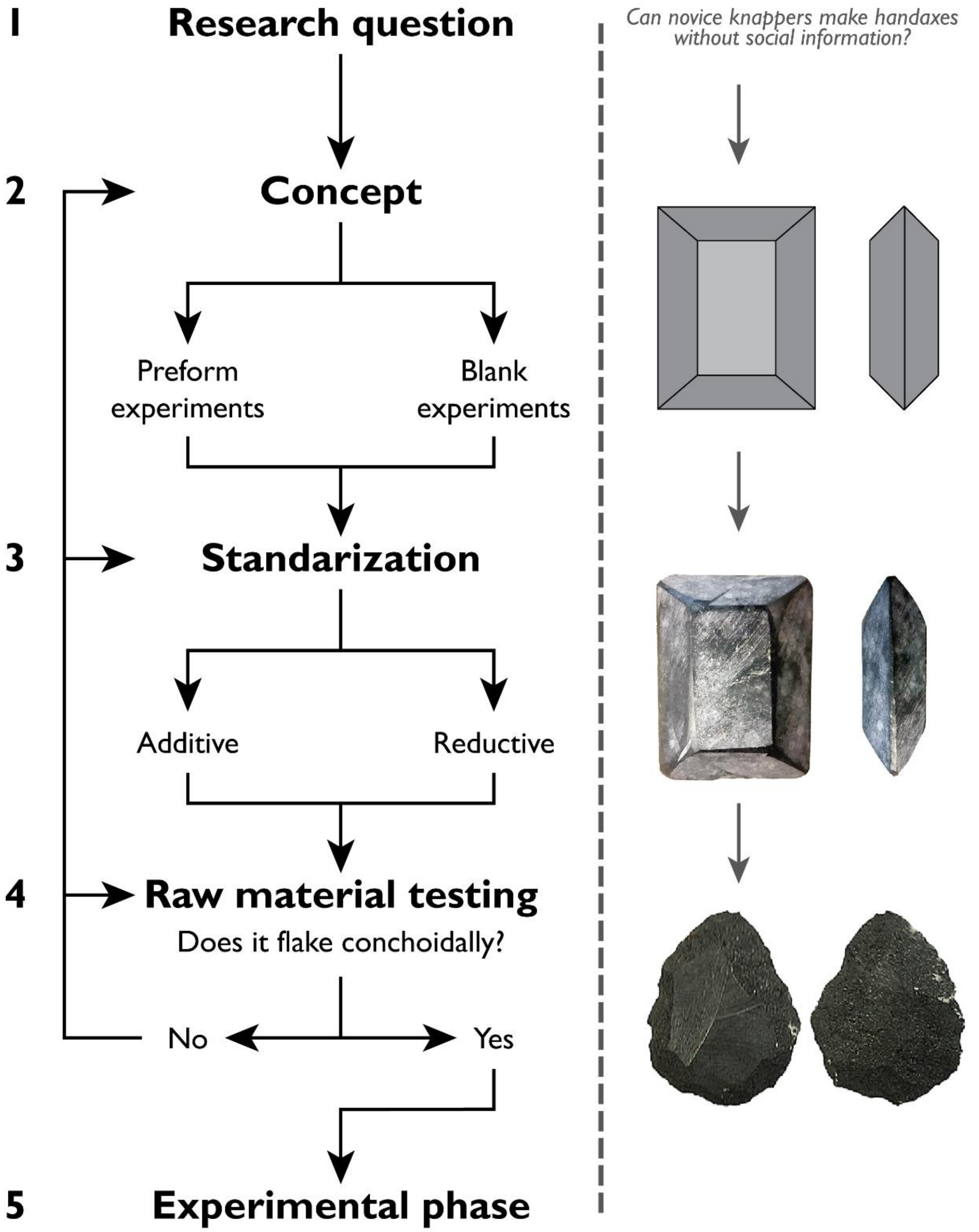


Figure 6 Stepwise program for the development of standardized blanks for knapping experiments. The concept for the knapping blank is determined by the research question, here with two main considerations: the use of preforms versus the use of blanks. The standardization step relates to the exact techniques used to make the standardized blanks, i.e., whether an additive technique (like porcelain-making) or a reductive technique (like a diamond saw and grinder) is used. Once the first blanks are created, they should be tested to see if they produce valid and/or desired outcomes. Only if this can be verified should proper experimentation begin. Images on the right are examples of this stepwise program in the specific case of sawed and grinded basalt blanks (see section 'Case Study 2: Machine-cut basalt' below'. Here, the protocol was shown to be feasible for generating knapping blanks for experiments on Acheulean biface production. Graphic by J.D.O-F, based on an earlier version by D.B.

tumblers. Automated machinery like milling machines would be even more reliable at producing standardized blanks (Lin et al., 2021). These techniques are the only way to standardize natural stones. Alternatively, stone can simply be used as it naturally occurs (with neither a reductive nor an additive standardization process). This might be considered still standardized in cases where erosive forces cause knappable blanks to be formed into similar sizes and shapes and also sorted into a substrate.

When developing standardized blanks (as in the case studies we present below), we generally followed a stepwise program (see Figure 6), beginning first with a specific research question in mind (i.e., what technology, behavior, or related concept is the focus). The research question determines the ideal form(s) that is to be used in the experiment. In some cases, this might only consist of rough dimensions or weights, but in much of our work, we used either pre-existing standardized 3D forms or designed idealized blank forms using Blender 2.8 (citation for the Blender 2.83 LTS manual). 3D models could then be printed out and used for the building of molds (e.g., for creating porcelain blanks). Once the standardization procedure has been carried out and the first blanks are available, it is necessary to test the blanks to determine where they are valid for the intended experiment. This includes determining whether the standardized blanks can be knapped at all (i.e., do they conchoidally fracture) and whether any desired target forms can be arrived at with the standardized blanks as a starting point (e.g., for studies on the Acheulean, can the blank be used to produce a biface). Since our particular focus is on Early Stone Age technologies, all of said raw material testing in the scope of the present text was performed with hard hammerstones, which were river cobbles from Germany (purchased at a building material supplier, precise provenance unknown). If and when the standardization procedure is reliable and the blanks are suitably valid, then the full series of blanks can be produced, and the study can enter the experimental phase.

Stepwise instructions on how to make standardized blanks and comparisons of different techniques and materials tend to be rare or incomplete in the literature. The aim of the present manuscript is to provide a more generalized guide to the practice of raw material standardization for controlled knapping experiments. We have divided up our efforts to produce standardized blanks into succinct case studies, with materials ordered from most to least valid. The subsequent sections are thus as follows:

- Hand-knapped flint,

- Machine-cut basalt,
- Manufactured glass,
- Porcelain slip and porcelain body,
- And Other materials (including sandstone, brick, Plaster of Paris, and craft concrete).

The specifics of the varied materials and standardization techniques are then followed by a broad overview that should be suitable as a quick reference guide (see Table 5). Ultimately, the presented options represent mainly what was available to use currently, but future developments in materials science and manufacturing technologies, along with the reduction in costs of relatively inaccessible materials and machinery, will open up new - and potentially even more valid and reliable - means of creating standardized knapping blanks.

Case Study 1: Hand-knapped flint

Flints and cherts have been a staple of experimental archaeology since its very origins as a practice (e.g., Johnson 1978), largely owing to their widespread abundance across the globe (Keller, 1981) and sharp, resistant edges that can be worked with greater ease than other materials, such as basalts and similar rock types (Luedtke, 1992). As with other kinds of stone, flints and cherts have been used extensively throughout history and prehistory, including during the Oldowan (e.g., at Bed II, Olduvai Gorge; Hay, 1976).

Though flint is a ubiquitous material, the quality of flint from source to source and even nodule to nodule can vary massively. Due to the nature of its petrogenesis, flint tends to be susceptible to irregularities like inclusions of non-silicified material, encrustations or cleft areas, which can all heavily affect the properties of flaking. Where good- to high-quality knapping material is required, the task becomes the identification of consistent sources of such flints or cherts and access to these sources. In our case, for instance, silica precipitates can be found across Germany, including our general region in Baden-Württemberg (Schürch et al., 2022), but local sources were judged to be generally low and/or inconsistent in quality and occur in such quantities to make material collection inefficient.

For our preparations, we identified and tested flints from two sources: pebbled nodules from Heidkate Beach near Kiel, Germany (hereafter referred to as Baltic flint) and nodules quarried from chalk deposits in Norfolk, East Anglia, UK (hereafter referred to as Norfolk flint; Figure 7).

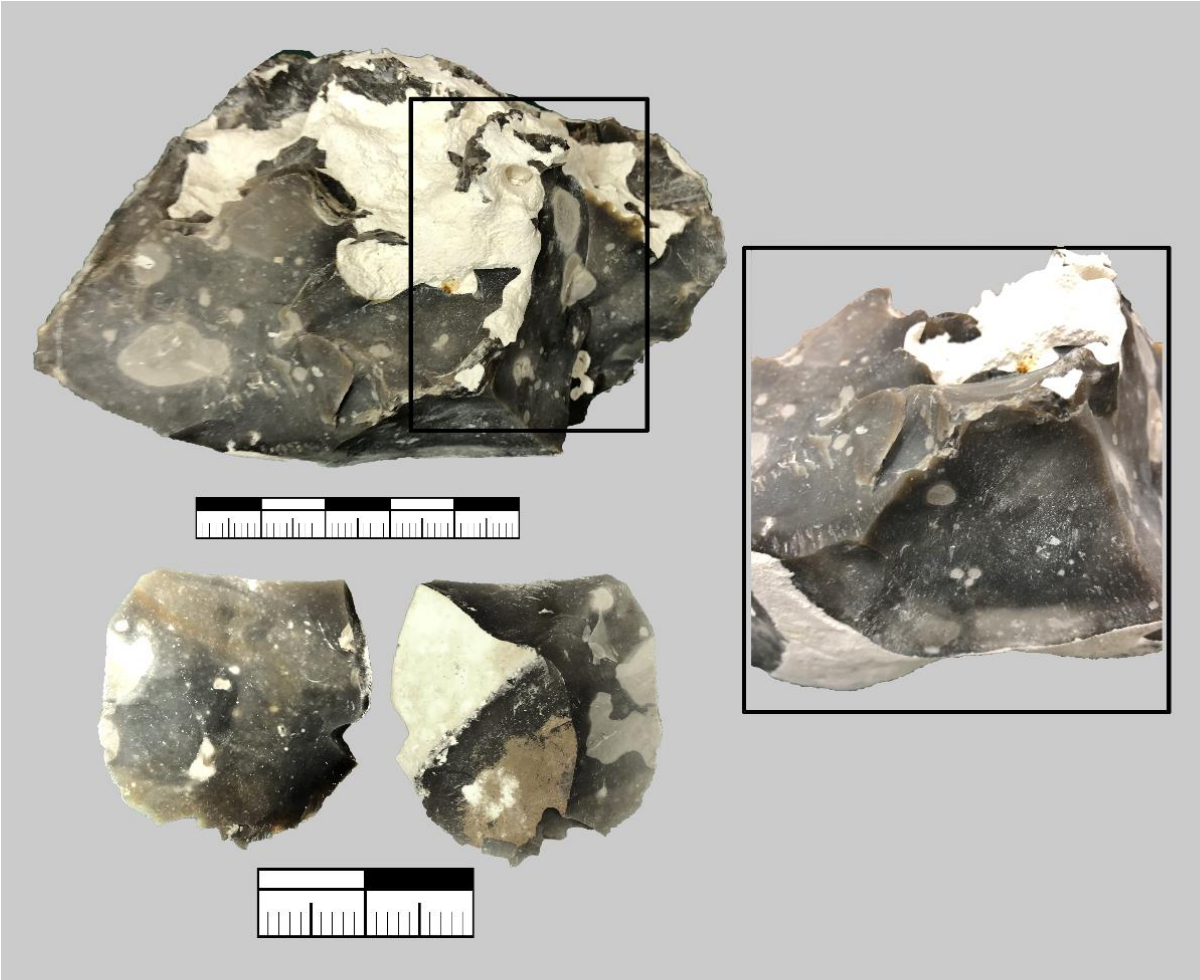


Figure 7 Flint from Norfolk, East Anglia, UK, that has been excavated from coastal chalk deposits. **Top left:** The full nodule. **Right:** A flake scar. **Bottom left:** dorsal and ventral perspectives of a flake. Graphic by D.B.

The Baltic flint was collected by one of the authors (DB) directly at the sedimentary deposits. At Heidkate Beach, the flint nodules can be found among other siliceous rocks, such as feldspar and quartz. Due to constant erosion by the Baltic Sea, the nodules are entirely rounded and occur in irregular round/oval shapes and in varying sizes, acting as a process of 'semi-standardization' (B. Morgan et al., 2015) and sorting. In the surveyed area, the flint nodules were fairly abundant, with about 10 nodules that were visually identifiable and easy-to-access within 15 minutes of strolling at the beach. In terms of quality, the Baltic flint was characterized by good conchoidal fracturing, which generated flakes with long-lasting sharp edges. The rounded outer surface and relatively thick encrustations that form the cortex, however, were a major barrier to initial fracturing, so that the nodules required additional force - when compared to non-encrusted flint nodules - in order to break them open. Laws and regulations regarding geological and archaeological heritage in the respective countries and localities should, of course, be considered before collecting and working with field-collected materials. Local authorities should particularly be informed, and the relevant permissions should be attained beforehand, in order to prevent infraction of legal frameworks and to prevent irrevocable destruction of natural ecospaces and endangerment of archaeological heritage. This may involve limitations on the quantity of permitted material collection, which is therefore a practical and logistical problem for large-scale studies, as well.

Unlike the wild collection of the Baltic flint, the Norfolk flint was imported via a stone-distributing company. The costs of the flint itself are relatively negligible; here, the main drawbacks involve the costs of transporting the material from the source to location of the experiments as well as the time required to transport it and the logistics of bringing it from site to site. The Norfolk flint exists in a variety of shapes and sizes as well as quality and is generally at least partially encased in chalk. In its raw form, the Norfolk flint is easier to knap, with the chalky surfaces being soft and unobstructive to removal and the exposed flint surfaces providing more easily worked knapping platforms than the rounded surfaces of the Baltic flint. For studies on the knapping abilities of non-human great apes (Bandini et al., 2021c; Motes-Rodrigo et al., 2022), as well as pre-study pilot experiments for a study on the same capacity in modern humans (Snyder et al., 2022), flake cores were prepared by simply knapping flint nodules with the use of a hammerstone (e.g., a roughly oval river cobble). For this purpose, the flint was worked by the knapper (DB) until it was approximately the mass and dimensions required by the experiment; the cores were also shaped to the extent that was achievable, given the quality and internal dynamics of the raw material. This further means that knapping platforms could not be purely standardized for their ease of flaking (hereafter also referred to as knappability). Knapping by hand involves an investment of both time and labor on the part of the responsible party (whether that is one of the experimenters or an otherwise affiliated expert) Also, the tendency of flint to vary in quality due to inclusions and processes of weathering and exposure means that not all nodules can be ‘molded’ into suitable experimental cores to distribute to subjects. Therefore, this particular method of standardization is neither efficient nor reliable, despite flints and cherts certainly being valid for the general purposes of replication experiments.

Case Study 2: Machine-cut basalt

Basalt is an abundantly occurring natural raw material (Yaroshevsky, 2006) that has been used for the production of stone tools at many archaeological sites and across geologic time (e.g., during the Oldowan of East Africa: e.g., Stout et al., 2005; Braun et al., 2009; Braun et

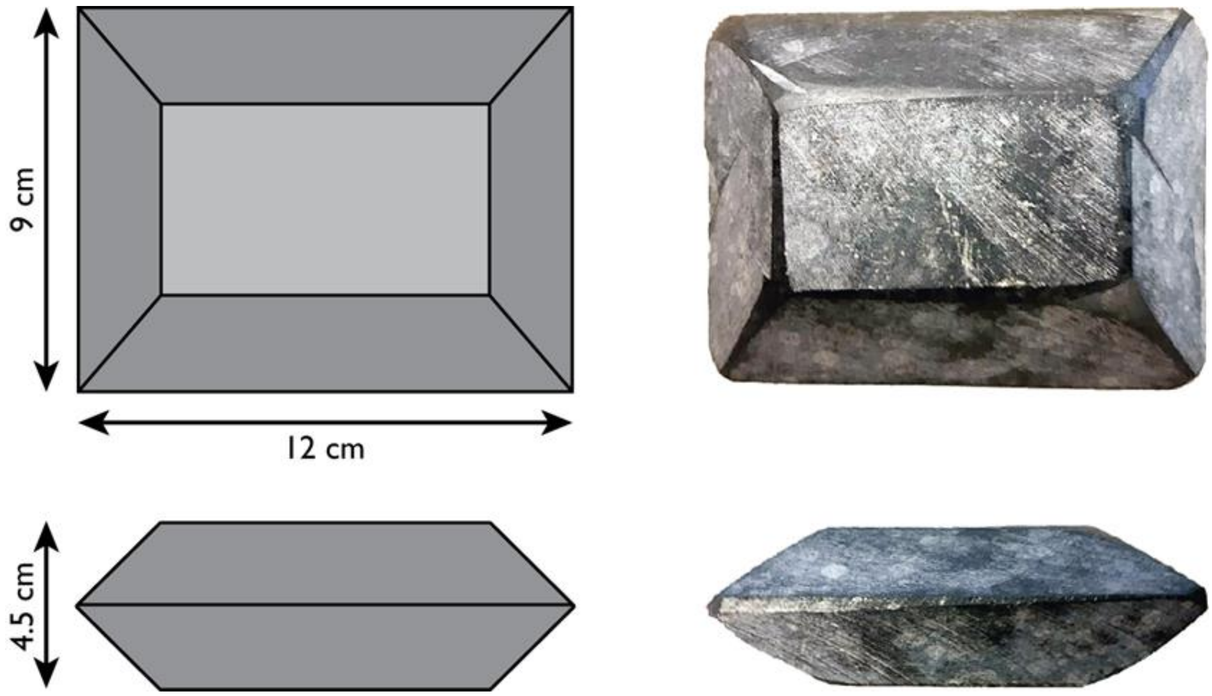


Figure 8 Schematized and actual polyhedral basalt blank made using a diamond-saw and grinder. Graphic by J.D.O-F, based on an earlier version by D.B.

al., 2019). The quality and knappability of basalts, as with other volcanic rocks, relates mainly to the size of grains, resulting from differences in the crystallization of the molten minerals during cooling and the conglomeration of different compounds (Militky & Kovacic, 1996; Stout et al., 2005; Braun et al., 2009). Fine-grained basalts are generally easier to work, while coarse-grained basalts are less suited for controlled flaking (Braun et al., 2009).

Basalt bricks were purchased from a local wholesaler of construction materials (Natursteine Park Tübingen) and had already been pre-shaped into rectangular prisms. The basalt was attributed to a source in the Czech Republic, most likely Vastany Hill. Relative to other raw materials, the basalt, in this form, was extremely cost-effective and easy-to-acquire, not just in one but many varieties. For instance, we performed principal knapping tests of a burgundy-colored, coarse-grained rhyolite in addition to the black/dark grey basalt that we decided to use in more extensive tests. As with other natural stone materials, basalt can vary within a single volume in terms of its density, homogeneity, and structural integrity (Klein & Langmuir, 1989; Farmer et al., 2003).

We selected bricks with the approximate dimensions of 25 cm x 10 cm x 10 cm, which seemed to be the best starting point in terms of yield, weight and general workability. To produce at least semi-standardized matrices (originally envisioned as an elongated pebble but

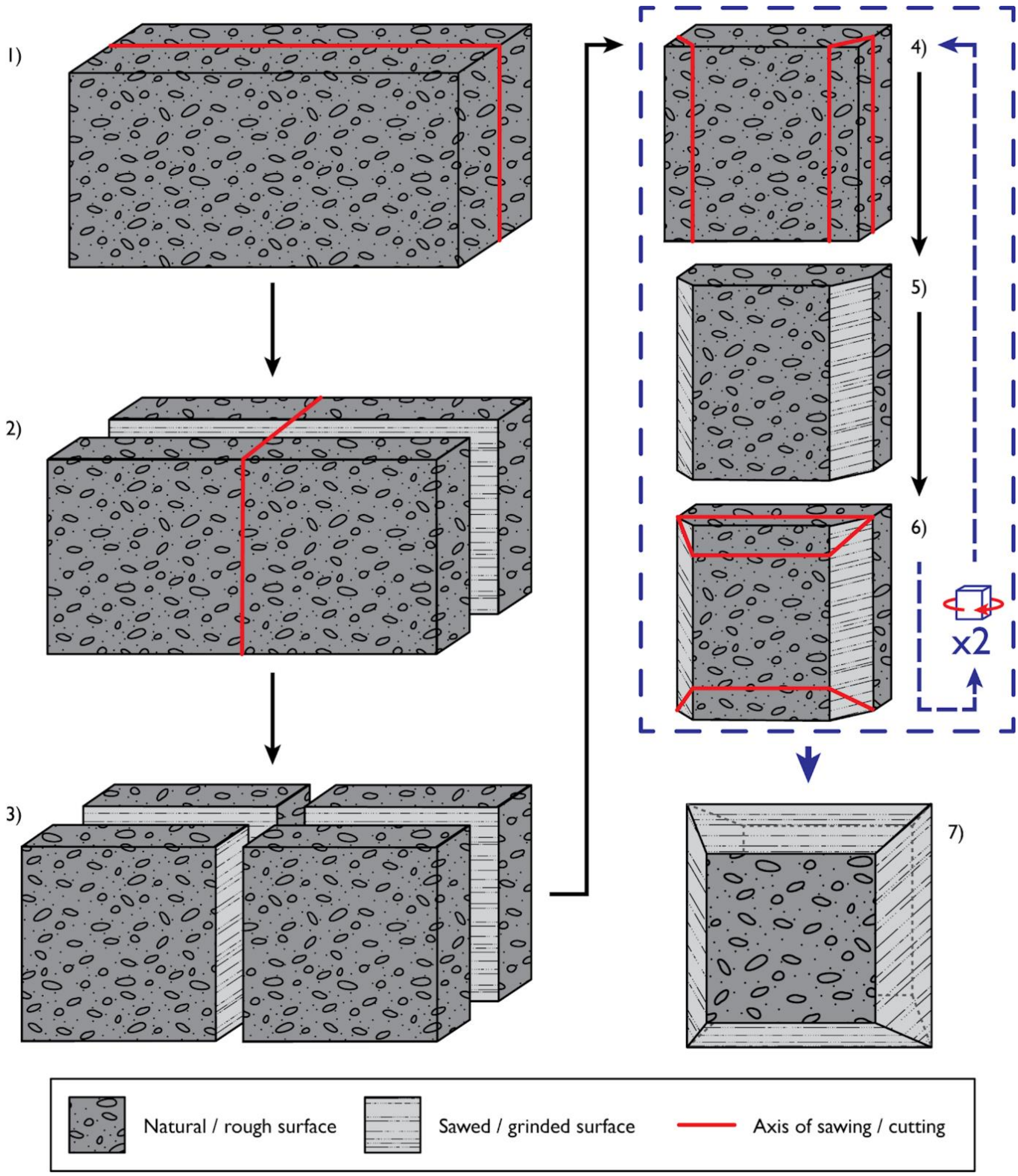


Figure 9 Visualization for protocol used to created standardized basalt blanks (and which would also apply for other types of stones). Descriptions of the relevant steps are presented below and on the following page. Graphic by J.D.O-F, based on an earlier version by D.B.

adapted into a polyhedral form as a result of the limitations of hand-guided machinery), the following series of steps was applied (Figures 8 and 9):

1. Estimation of the most regular face and further flattening with a diamond-band-grinder, designed and usually used for glass manufacturing (Glaskant-S, Knopp Maschinen GmbH),

2. Cutting the brick in half along the longitudinal axis with the use of a diamond-blade saw, again designed and usually used for glass manufacturing (Sägboy I by Knopp Maschinen GmbH; saw blade made of a steel-copper alloy and coated in diamond dust),
3. Alignment along the created flat surfaces (~ 25cm x 10cm x 5cm) to saw and grind the irregular, lateral surfaces, resulting in a semi-regular cuboid,
4. Sawing into two halves (each ~ 12.5cm x 10cm x 5cm) of the cuboidal form,
5. Fixation of the cuboids into a wooden frame, angled at 45°, and sawing of the angled blank to create a 45° edge around the corners and up to the medial axis of the 5cm lateral surface (done to all sides and from both surfaces), and
6. Fine finishing of the object by grinding the rough edges and rounding the corners into a more regular shape, as in the exemplary depicted matrix with dimensions of 12cm x 9cm x 4.5cm (deviations from these approximate dimensions being the result of the thickness of the saw blade and the material lost while grinding).

The described standardization protocol of shaping basalt by hand is more widely applicable to other naturally occurring stone materials, as well as to glass. However, the ease of standardization, efficiency of the process, and the stringency of safety protocols need to be attuned according to the differing characteristics and qualities of the material. Here, we describe mainly the process for basalt, though glass was also shaped in this way during our standardization trials. Security precautions are an essential element of the protocol; it is necessary to constantly wear gear such as safety glasses, gloves, ear protection, and an apron while sawing and grinding the basalt (or other material). Due to the constant water flow of the blades' cooling and dust-bending system, closed-toed shoes with well-profiled soles are recommended to prevent the processor from slipping and in case quick reactions are required. If possible, we recommend a separate, enclosed, and tiled room for housing the machinery. Furthermore, the right machinery and respective attachments are pivotal to successful processing. During the process of cutting basalt, different types of grinding bands and sawblades were tested and even partially destroyed. Especially while cutting the entire brick in half (as in Step 2 described above), the saw was likely to become extremely tense and get stuck due to the heating and slight expansion of the sawblade, in combination with momentum and vibration caused by the motor and sawing movement itself. The best results were achieved when moving the sledge slowly and allowing the saw to "find its way" through the material in a gentle, guiding fashion. Using this technique, the saw needed between two and four minutes for a single cut. As another side effect, material from the saw blade can potentially adhere to the stone, thereby modifying the material properties. If tension (see above) built up while cutting near the edges of the mass, the material sometimes broke apart and was ejected outwards like a projectile.

Even with pre-shaped knapping platforms as generated by our sawing and grinding protocol, basalt is still very tough and therefore requires more force from the knapper in order to produce suitable flaking outcomes when compared to flints and cherts, for example. When knapping basalt with larger grains, the resultant flakes can have somewhat jagged edges (rather than smooth cutting edges like from finer-grained basalts or from silica precipitates). Nonetheless, knapping basalt (Figure 10) ultimately results in usable cutting edges that are definitely viable for extractive foraging tasks, whether real or simulated. In the case of basalt blanks, we performed additional tests with novice individuals (i.e., archaeological and primatological researchers with little to no hands-on knapping experience), reaffirming the prior observation about basalt's properties. Given the difficulty involved in achieving controlled flaking of the material, basalt is at least not an ideal material for teaching contexts or for experiments involving inexperienced or novice toolmakers (unless otherwise prescribed by the need for external validity).

Figure 10 A 'handaxe' produced by knapping one of the machine-cut basalt blanks, demonstrating the validity of the blanks for experiments related to Acheulean toolmaking. Graphic by J.D.O-F, based on an earlier version by D.B.

Case Study 3: Manufactured Glass

In the archaeological record, there has been extensive use of obsidian and other naturally occurring volcanic glasses as knapping material by modern humans and pre-modern hominins in various locales around the world, starting at least 1.7 Ma (Piperno et al., 2009). Through colonial interactions, manufactured glass was introduced to societies around the world and subsequently adopted (independently) by many groups as a raw material for knapping traditions (i.e., remanufactured glass industries, e.g., Cooper and Bowdler, 1998). Due to its model properties of conchoidal fracture, glass has been an important and reliable material for controlled experiments in archaeology (Dibble & Whittaker, 1981; Dibble & Pelcin, 1995; Dibble & Režek, 2009; Dogandžić et al., 2020). Using soda-lime glass, experimenters have previously been able to control for extraneous variables in order to study what determines the characteristics of flakes. Glass is especially suited for controlled experiments due to the reproducibility of the same sized and shaped glass blanks, an advantage not shared with all materials (Dogandžić et al., 2020).

While natural glass like obsidian is restricted geographically and available only from raw material suppliers in countries with the right geological context, manufactured glass is extensively available in industrialized nations. Due to the specific context of our location in Germany, we did not have access to a cost-effective and logistically sound source of obsidian, though this constraint would not be relevant for all researchers (such as in North America where obsidian is reasonably cheap and ready-to-acquire).

Pre-made glass forms can be purchased from online retailers and glass manufacturers and wholesalers. These glass forms can be either clear or colored with additives, which should not affect the fracturing properties of the glass itself. Glass forms can also be made-to-order; we ourselves had contact with multiple companies offering this service, but an order of glass blanks based on a 3D model of our own design ultimately did not come to fruition. The process of molding solid glass into the desired form requires high temperatures and can be extremely volatile, and in our case, resulted in the mold breaking apart, making it impossible to produce the desired elongated pebble forms. Other manufacturers did not have the capacity to produce the forms we required in general (e.g., completely solid without a hollow space inside, molds large and stable enough to produce the required volume for the objects). Therefore, it is pertinent to ensure beforehand that the manufacturers possess the

Figure 11 Manufactured glass hemispheres, including 'black' glass **(A)** and spray-painted clear glass **(B)** hemispheres and flakes made from each (center left and center right). Graphic by J.D.O-F.

infrastructure necessary to create the commissioned blanks. A separate alternative to consider is self-made glass blanks, which can conceivably be created at lower temperatures by melting together recycled glass pieces, though this does not allow for the same control over material quality and homogeneity as with freshly formed glass at higher temperatures. A large 'nodule' of recycled glass was provided to one of the authors (W.D.S.) but did not function as a good knapping material because of the combination of various glass types that generated internal fault lines and prevented consistent, controlled fracturing of the object. As with pre-made glass, custom-made can also be colored with additives.

Though we did not work with obsidian or any other naturally occurring volcanic glass in our efforts, there is previous documentation of the standardizability of this kind of material. Sheets & Muto (1972) used a method not unlike the one we applied to basalt in order to create blanks for blade-making. We also used the method described above for creating polyhedral basalt blanks to generate aesthetically similar standardized blanks from colored glass bricks (rectangular prisms, ~10cm x 10cm x 3cm). This glass appears to be black in color from the outside but is actually revealed to be a deep purple upon the removal of flakes. Coloring of the glass creates at least a superficial resemblance to naturally occurring volcanic glasses. Additional grinding of the glass surface can be used to produce a rougher cortex not dissimilar to what can be found on volcanic glasses as well.

Ultimately, we ordered a large series of pre-made clear glass paperweights (hemispheres by design, diameter: 10cm, height: 4cm; Figure 11) to be used mainly as “split cobble” blanks for least-effort flaking experiments (Snyder et al., 2022). These glass hemispheres were sprayed with a light grey spray paint in order to generate a sort of pseudo-cortex. Only one layer was required to cover and adhere to the glass forms and was applied in two sittings to ensure full coverage; the round side was sprayed first, allowed to dry for a few hours (or as needed depending on the temperature and humidity of the surrounding air), and then the flat side was sprayed and left to dry again. A fraction of the half-spheres required repainting due to immediate chipping of the paint or light scratching, but otherwise the process was relatively seamless. The creation of the cortex on glass, whether via grinding or painting, has multiple benefits, as it can create a more naturalistic appearance of the material (as far as aesthetics might be a relevant dimension in the experiment), and it facilitates attribute analysis and refitting of artefacts *post hoc*. Still, reflectivity of glass, clear or colored, can be problematic for imaging and analysis (e.g., when using a visual light scanner, it is necessary to spray glass artifacts with temporary anti-reflective spray in order to capture the details of the object surfaces).

As expected, and in all cases, the glass was extremely easy – and consistent - to flake and effective for use in cutting tasks due to the bountiful sharp and acute cutting edge produced via knapping. The layer of spray-paint (about 1 to 3 coatings, in order to completely cover the round glass surface) seemingly does not impact the knappability of the glass surface - whether in a positive or negative fashion.

Case Study 4: Porcelain slip and porcelain body

The history of ceramics being used in flint-knapping experiments dates back ostensibly several decades (Johnson, 1978). Recently, increased attention has been paid to porcelain as a suitable alternative to stone that can be quite easily standardized for controlled experimentation (Khreisheh et al., 2013; Ranhorn, 2017; Stade, 2017; Speer, 2018).

Powdered clay and moldable porcelain were ordered online from a local (i.e., German) ceramics distributor, and then processed internally by one of the authors (WDS) for the creation of standardized forms. Material costs for porcelain production are relatively low, as reasonably priced porcelain clays and bodies, as well as plaster of Paris for molding and basic mixing tools and containers, are fairly easily accessible. We also pursued the option of 3D-printed ceramics, but this avenue was abandoned for several pertinent reasons: individual blanks would be extremely costly (quoted as >$100 per blank), 3D printed blanks would need to be hollow, and the firing temperature was well below that required for the correct material properties (800°C) (see also Discussion below).

We selected porcelain that could be heated in the range (Cone 6) achieved previously by other researchers (Khreisheh et al., 2013; Ranhorn, 2017; Stade, 2017; Speer, 2018). For early trials, we used porcelain clays that required the addition of water to form a slip. Mixing of the clay powder into the water (approximately one part water to 2-3 parts powdered clay) for the formation of the slip lasted approximately one hour, after which the slip was allowed to mature for one to two days (Walker Ceramics). The slip was then poured into the predetermined plaster mold (i.e., with forms based on cereal bowls and 3D-printed Dibble cores) until the slip was slightly overflowing. At this point, the porcelain slip was allowed to sit, during which time a loss in volume was incurred. Volume loss is related to the shape and the size of the desired form, due to the relationship between water loss and retention and surface area (citation). In the case of a single-piece mold and slip-form clays, the drying of the slip results in lipping on the upper edges, which can be avoided by modification to the mold design or removal of lipped portion with piano wire or some other tool.

Solid porcelain bodies can also be used instead of a slip formula. Porcelain body only needs to be pressed into the plaster mold and allowed to solidify at room temperature (Figure 12), thus not resulting in the same extreme volume loss to the prepared blank. The pressing of the porcelain body, however, has the disadvantage that internal fissures can form, disturbing the homogeneity of the object and creating potential for pieces to explode in the firing phase.

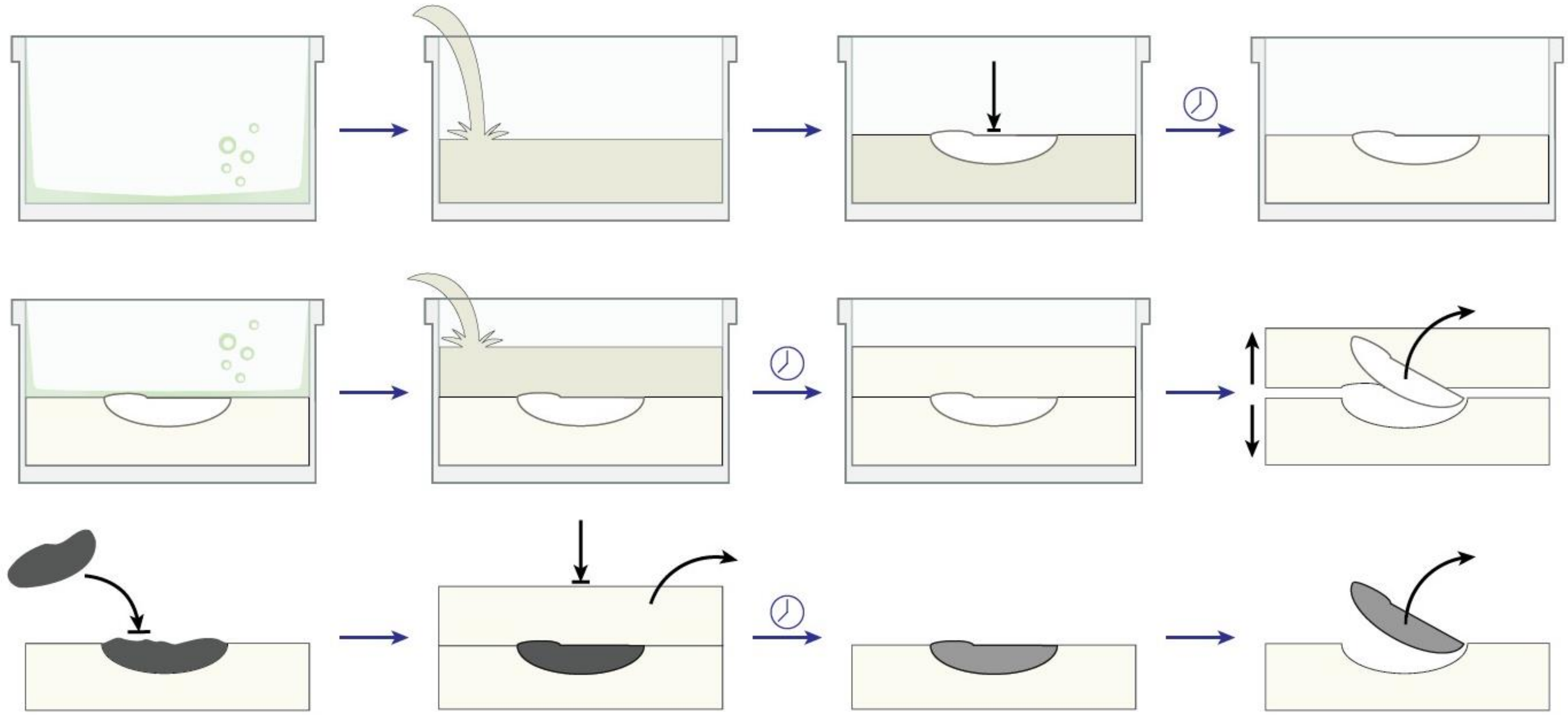


Figure 12 General procedure (as described in the text, see preceding and following pages) for production of porcelain blanks (pre-firing). Graphic by J.D.O-F. based on an earlier version by W.D.S.

Figure 13 Unfired porcelain blanks. Graphic by W.D.S.

Again, this can be evaded by using more material than necessary when pressing, as using just enough encourages deformation of the material in such a way to form fissures. The excess porcelain body can then be removed with a tool or wiring; just by circumstance, it was discovered that removing excess body with gardening wire can create impressions in the porcelain that mimic the appearance of ripples as with natural flakes (Figure 13).

After at least two days of settling, the porcelain blanks were then fired in a kiln (model of kiln here; Laboratory of Soil Science and Geoecology, University of Tübingen). If the blanks are not allowed to settle for enough time before firing, then water molecules will not have fully escaped, and the blanks might either explode in the kiln or simply develop fissures and air pockets that reduce knapping quality. The same may be true if the blanks contain any fissures that form during the process of pressing them into form (as with porcelain bodies) or if the blanks are heated for too long in the kiln (e.g. one batch of porcelain that WDS prepared exploded after being kept in the kiln for approximately 48 hours, as overnight ambient temperatures were too high to allow for safe removal). The pieces of porcelain were heated to a maximum heat of 1240 degrees Celsius (Figure 14) and then left to cool overnight. For the white porcelain slip, the blanks were heated to 1240 degrees Celsius in three hours, held at this temperature for three further hours, and left to cool overnight. For the black porcelain body, the heat was increased for nine hours, held at the maximum temperature for one hour, and subsequently left to cool. These heating curves were determined to be effective for reliably producing blanks with excellent knapping qualities (though refer back to prior mention of potential issues). The firing method for porcelain slip is also more expedient than previously reported methods (cf. Page, 2014) while samples of porcelain body required more prolonged heating in order to prevent fracturing of the material. As noted elsewhere (Page, 2014), we observed conspicuous volume loss due to firing.

Though porcelain blanks are highly standardized, the process required to produce them is inefficient. Given enough molds and clay, the blanks can be produced in large quantities, but this still necessitates relatively lengthy periods for slip formation, drying and heating (mold creation can also be time-consuming). The largest bottleneck in the production process is at the firing stage. Space inside the oven naturally determines how many blanks can be fired at

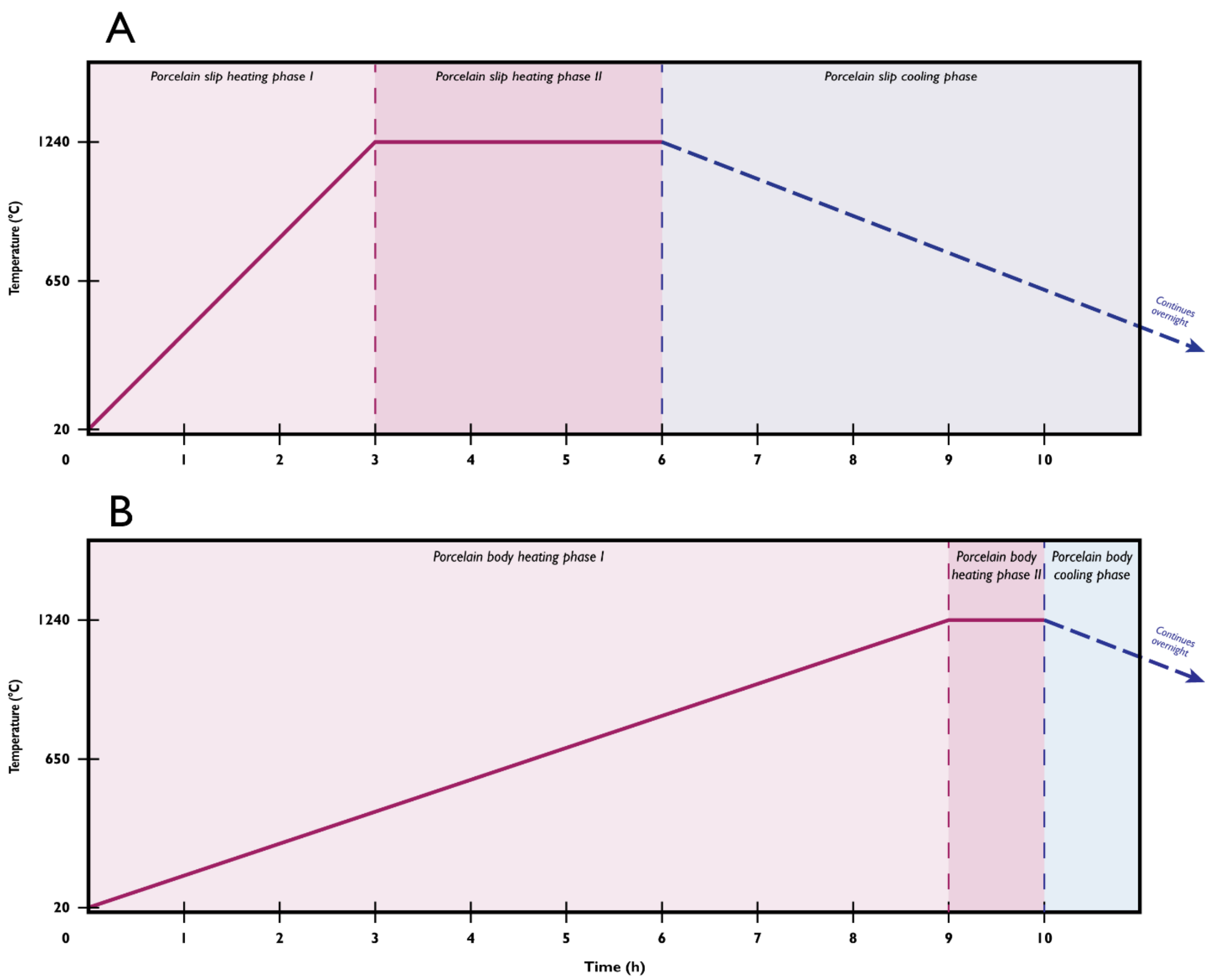


Figure 14 Heating curves for the firing of porcelain blanks, including the curve used for white porcelain slip **(A)** and black porcelain body **(B)**. Graphic by J.D.O-F. based on an earlier version by W.D.S.

once, and placement inside the oven at once can lead to quality differences between blanks. Also, there is no guarantee that the blanks will have the required properties in the end, as errors due to drying or firing may not be immediately apparent. The likelihood of low-quality blanks reduces, however, with practice and greater adherence to a careful and meticulous production process. All steps require attentiveness and an intensive investment of labor by the experimenter (or whomever is preparing the blanks).

Porcelain has physical properties on par with flints and cherts typically used for knapping (Khreisheh et al., 2013). Porcelain cores readily produce sharp, usable flakes, with very little production of dust and small shatter. Porcelain cores are also lighter than stone or glass cores of similar size, which would make them safer and more accessible for use in novice knapper experiments as well as in teaching or learning contexts. Visually, standard white porcelains can be very similar in hue to light-colored cherts; it is also possible to color porcelain or to buy pre-colored clays and bodies. We purchased, for example, a black-colored porcelain body that appears superficially like volcanic materials (e.g., basalt) after being fired. We also observed that - upon firing - the outer layer of porcelain becomes molten, essentially re-creating a cortical covering of the material. When heated longer, not only do the blanks have a lower likelihood of exploding, but also the resultant cores knap more like, e.g., a fine-grained basalt than like a higher-quality flint, as was the case with the relatively fast-heated white porcelain slip.

Other Materials

We also performed smaller-scaled trials of other materials - with and without explicit standardization protocols.

As with the basalt, bricks cut and/or broken from naturally occurring sandstone were purchased from a local construction wholesaler. Again, these were roughly in the form of slabs or rectangular prisms. Sandstone was certainly utilized for tool production in the Paleolithic (e.g., McNabb et al., 2004; Kuman et al., 2005; Hernandez et al., 2012), but for our purposes was determined to be too tough to provision to human novices. The same protocol and considerations that applied for machine-cut basalt would still be viable if one were to opt for sandstone in controlled experiments. As another alternative, 3D-printed “sandstone” was also tested (similar to 3D-printed ceramic but with a distinct texture and hardness). Two samples were ordered from an online 3D-printing retailer but were not found to conchoidally fracture in the desired manner (the samples simply snapped when struck). Further unfavorable attributes were the limits of blank size/volume (mainly due to the limits of what is still a rather novel technology) and the excessive expense of the 3D-printed sandstone with respect to the actual volume achieved. Therefore, it was not possible to produce material in either the right size of individual blanks or cumulative quantity needed for experimentation purposes.

In addition to sandstone, we also tested the properties of various mass-produced construction materials. Masonry bricks - for instance - are one human-made raw material that already has an established usage in archaeological experiments (Geribàs et al., 2010; Lombao et al., 2017). Bricks have been described as having the “same mechanical properties (conchoidal fracture)” as natural stone, while being homogenous, standardizable, and safe for novice knappers to use (Geribàs et al., 2010). Bricks were purchased from a local construction and hardware store (TOOM Baumarkt, Germany). The bricks were selected on the basis of the maximum available continuous volume that can be exploited. Bricks which, due to weight and material savings, have cavities or slots are problematic, as these pre-determine and limit the breaking. The main advantage of bricks is that they can already be purchased in a standardized format, and their ubiquity further means that an enormous variety of bricks is available for selection. Bricks are still quite hard, and their shapes are typically not suitable for knapping platforms (though alternative shapes with better exterior angles are available or could be made to order), so controlled flaking is not particularly easy when using bricks. They are also often stored outdoors, which can result in reduction in the quality of the material and its flaking properties.

Not unlike bricks, concrete is another human-made raw material that can be conchoidally flaked in the right circumstances. Here, we tested mainly case-burned limestone, which is generally a simple mixture of burned limestone with plaster and other elements in different ratios. Given its abundance and cost-effectiveness, we also tested concrete for its appropriateness for knapping experiments, finding that it can indeed be conchoidally flaked and also can produce effective cutting edges. Similar to the bricks, concrete was purchased from a construction supplier and is generally available at most hardware stores or stone traders. We purchased a series of concretes with different compositions in order to assess a range of possibilities. In every case, these concretes performed fairly similarly in terms of knappability. Though for our short trials, the concrete was already pre-made in standardized brick forms, the usual process of concrete production means that it can be made to many potential desired forms, especially those that would offer better knapping platforms than a simple rectangular prism does.

Plaster of Paris, also known as gypsum plaster, is traditionally a material used in construction or decoration - or in the case of the aforementioned porcelain-making, molding). As with porcelain slips and bodies, plaster of Paris in its powdered form was purchased from an online ceramics supplier. The powder was mixed with water (a standard ratio of two parts of powder to one part water), a desired form was fixed to the bottom of a container (e.g., cereal bowl or 3d print), and then the liquid plaster mixture was poured into the container until the positive was sufficiently covered. After a few hours of sitting, plaster is usually cold and solid to the touch, indicating that the mold is finished and can be removed from the frame. Opportunistically, we attempted to fracture plaster molds that had failed or were no longer needed. When struck with a hard hammer, the plaster fractured conchoidally, resulting in flakes with acute angles, but which were not particularly sharp and therefore not useful as actual cutting tools. For strict replication regimens, plaster is not really viable, but for the purposes of education (cf. Shea, 2015; Clarkson, 2017), plaster services as an easy-to-make and especially easily standardized knapping material.

In addition to the preceding materials, we tested two more that did not conchoidally fracture. The first was a craft concrete (Viva Decor; containing Portland cement) that - when mixed with water - can be kneaded by hand or pressed into a mold and then air-dried. Though the craft concrete does fracture when struck with a hammer, it contains filamentous structures that hinder the formation of fractures and especially the detachment of flakes from the blank. The detachments intermittently resemble conchoidal flakes but lack any usable cutting edges (again, due to the structure underlying the craft concrete). Thicker forms fracture better than thinner ones, as far as conchoidality. The craft concrete is effectively inappropriate for actual knapping studies, but would be reasonable for demonstration purposes (e.g., for public outreach, at schools, again cf. Clarkson, 2017; Shea, 2015).

Acrylic resin was also briefly tested, given that it can be filled into a mold and allowed to set and dry at room temperature without any additional steps. The texture and fracturing properties of the acrylic resin were fairly similar to something like a bar of soap, which would indicate that it is not a valid material for any of the contexts we have underlined.

Table 5 General overview of all materials we tested, including their validity, reliability of standardization techniques, costs, and recommendations.

Material	Validity		Standard. Technique		Costs	Recommendation
	Archaeological	Knappable	Method	Reliability		
Basalt	Yes	Yes	Knapping	Low	Cheap when purchased in wholesale, potentially high labor costs, dangerous due to material toughness	Valid for studies of Early Stone Age, semi-reliable, low costs but safety concerns
Flint, chert, etc.	Yes	Yes	Tumbling	Low	Cheap when purchased in wholesale, potentially high labor and transport costs	A near-universal standard for knapping experiments, valid and semi-reliable
Sandstone	Yes	Yes	Saw / grinder	Medium	Cheap when purchased in wholesale, potentially high labor costs, dangerous due to material toughness	Valid and semi-reliable, but not especially easy to knap or standardize so it is not highly recommended
Granite	Yes	Yes	*For all stone types.		Cheap when purchased in wholesale, potentially high labor costs, dangerous due to material toughness	Viable as anvils but otherwise not particularly useful due to difficulty to knap and standardize
Glass	Yes	Yes	Saw / grinder	Medium	Readily available pre-manufactured (incl. online) and can be made-to-order	Highly recommended due to reliability of standardized forms and known knappability
			Manufacture	High		
Porcelain	Yes	Yes	Slip pouring	High	Cheap, but the process is time-consuming and inefficient and requires higher labor costs	Highly recommended due to reliability of making standardized forms, but with the caveat of the time- and labor-intensive processing
			Body pressing	High		
Brick	No	Yes	Knapping	Low	Cheap and easily accessible, time and labor costs of manufacturing custom blanks not determined by the authors	A handy solution for experimental blanks due to prices and availability, but knappability is not so ideal
			Saw / grinder	Medium		
			Manufacture	High		
Concrete	No	Yes	Saw / grinder	Medium	Cheap and can be poured to form by hand, also not especially time-consuming	A decent material for knapping experiments, due to the ease and low costs of standardization, but material is tough to knap
			Manufacture	High		
Plaster of Paris	No	Yes	Manufacture	High	Cheap and can be poured to form by hand, the process is extremely quick and requires little labor input	Potentially useful for demonstration purposes or experiments, so long as sharp cutting edges are not desired
Craft concrete	No	Yes	Manufacture	High	Can be quickly and easily produced, but high relative costs	Filaments in craft concrete make it a poor material for knapping experiments, but it could still be usable for demonstrations
Acrylic resin	No	No	Manufacture	High	Reasonably cheap and not particularly time-intensive	Absolutely unsuitable for use in knapping experiments or demonstrations

Discussion

General overview of materials

In the preceding text, we have presented the results of various standardization attempts and raw material tests that we performed over a multi-month period. This procedure inspired us to report our findings and observations so that those interested in the possibilities of raw material standardization would have access to replicable protocols and - at least some basic - comparisons of various potential techniques (see Table 5).

Validity

External validity is the justification for generalizing that extend beyond the context of the study, e.g., via analogy between experimental processes and archaeological processes (Mesoudi, 2011; Lycett & Eren, 2013; Eren et al., 2016; Pargeter et al., 2019). Maximal external validity in archaeological experiments exists when the experiment uses the same precise raw materials as available to the hominins that researchers wish to study. Otherwise, the next niveau of external validity would be those naturally occurring stones, minerals, and glasses that are known to have been utilized in prehistory. At minimum - and as we present here - materials should be knappable to be externally valid, even if these materials do not appear in the archaeological record and/or would not have been available to stone toolmaking hominins. We therefore exclude non-knappable materials (e.g., Styrofoam or plasticene as in e.g., Schillinger, 2014; Lycett et al., 2015; potatoes as in Clarkson, 2017; or acrylic resin as described above) in terms of validity as they therefore require separate behavioral or action patterns to be modified or shaped and consequently lack any ecological validity (i.e., producibility of sharp tools for extractive foraging; Toth, 1985; Wynn and McGrew, 1989).

Knappability (otherwise, "ease of working, e.g., Khreisheh et al., 2013) is an important consideration when selecting raw materials for knapping studies (especially those with novices) and for teaching purposes (including for public outreach). The grading system – here adapted from Khreisheh et al. (2013), which was there itself adapted from Whittaker (1994) – serves as a useful indication of how workable certain materials are relative to other knappable materials. We have further built upon this by grading the artificial materials that we tested, and which have yet to be incorporated into the scheme (Figure 15 and Table 5); we also graded porcelain to see if the values reported by Khreisheh et al. (2013) could be replicated.

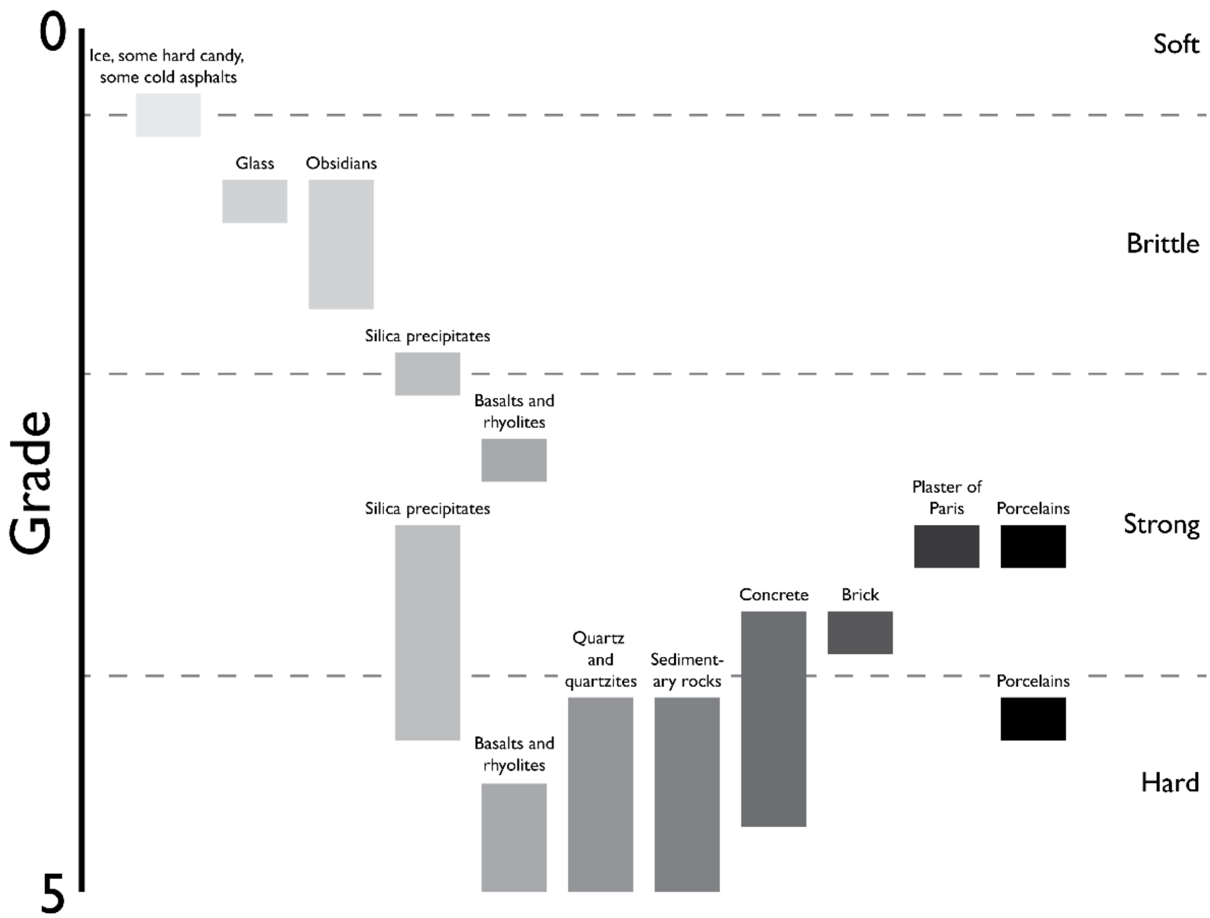


Figure 15 Gradient of knappability of raw materials from softest (0) to hardest (5).

Grading of knappability was performed by both first authors (D.B. and W.D.S.; each has previous experience knapping many of the pre-graded materials) for reliability. As previously reported, porcelains were similar in quality to finer flints and cherts. Plaster of Paris knaps about as well as either flints or cherts, as well, but the material is somewhat chalky and knapping of it does not readily produce usable cutting edges. Plaster of Paris and craft concrete would then be useful where external validity does not need to be maximized, i.e., for teaching and public outreach (cf. Shea, 2015; Clarkson, 2017). The various types of concrete we tested were fairly similar to tougher natural raw materials like quartzites, rhyolites, and basalts, while brick was closer to coarser flints and cherts.

Reliability

The relative ease, efficiency, and consistency of standardization varies widely between techniques and materials. Many different factors play a role in determining just how reliable standardization protocols for a material are, and therefore, many different factors must be considered in the process of designing knapping experiments that involve standardized blanks. Logistics is one such factor to consider. Although some materials might otherwise be valid and reliable, it may not be feasible to use them for experiments, for such reasons as transportation (e.g., moving natural stone from the source to the testing sites/facilities), labor investment whether that means hand-based labor or individual expertise, time and availability of equipment for processing, and so on. Equipment and machinery are major obstacles, since devices like diamond saws, rock tumblers and milling machines can be expensive and require transport costs, space to be installed, regular upkeep, and energetic demands like electricity. Certain technologies are only available as affordable to researchers from a limited set of manufacturers, and these manufacturers are not necessarily present and accessible in the home countries of the researchers. Material quantity and quality is the next major consideration. Essentially, one must determine not only if the quality of the material is suitable and valid, but if enough units of blanks can be ordered or made consistently to the necessary standard. Ordering natural stone in bulk, for example, requires the additional acknowledgement that the material quality will vary, which would often necessitate that more material should be attained than is actually needed for implementation in a study. Even beyond this, samples should also ideally be attained and tested beforehand to ensure the suitability for usage in research or teaching, so that large amounts - or really any at all - of insufficient quality, impractical or unusable material is not inadvertently stockpiled. Another part of this equation is the steepness of learning curves for standardization techniques. In cases where someone involved is already very familiar with the technique, this problem is pretty negligible, but if the process is new to the experimenter(s), sometimes this means an adjustment period for learning how to actually apply the technique for making blanks. Learning to perform the techniques therefore involves costs pertaining to investment of time, labor, and materials for practice. Given the wide range of factors that affect the reliability of materials and techniques (those listed here being just a handful of examples, really), the individual factors should not be considered alone, but also with the trade-offs of different variables balanced against each other. Reliability is in essence that - the balance of all variables related to easy, efficient and consistent production of standardized blanks.

Untested techniques and future prospects

In our preparations, we were not able to exhaustively test all available or potential techniques and materials. We considered, for example, the possibility of using a rock tumbler to generate polished - albeit not truly standardized - stone pebble blanks; gardening and construction companies already do this in order to create pebble-like forms for decorative purposes, though generally these fall below the size range useful for proper knapping experiments. The rock tumbling mechanism for our purposes is instead intended to roughly replicate the process of erosion that produces pebble forms in rivers but in only a fraction of the time. Other possibilities would include the pouring of concrete into a specified mold, pouring of molten glass into a specified mold (the obstacle for us being the purchase of or access to an oven that can reach the required melting temperatures for glass production), and the cutting of rougher materials like basalt or granite into shape using a water jet rather than a diamond saw (we had contact with at least one company capable of performing this).

Automated milling machines may be the best option overall for consistent production of uniform blanks across material types, stone included. These can be programmed to cut material according to a preconceived 3D design, but individual blanks can require inordinate amounts of time to be shaped to the design and the technology itself is not particularly affordable (Lin et al., 2021). Financial concerns and the need to import the machine from abroad were the main reasons why we were unable to acquire a milling machine; hopefully, more companies will be able to produce the proprietary technology and affordably so, so that experimental archaeologists can make wider application of the automated milling machine's potential.

3D printing is one avenue that we pursued for our research projects. Presently, commercial printing companies offer printing of objects in glass, porcelain, and sandstone, along with many other materials. Given these are relatively new technological developments, the processes and the products thereof are still being refined. Indeed, the process for printing porcelain, so far as we were informed, does not meet the requirements for the production of suitably knappable material. The maximum firing temperature for 3D-printed ceramics - as we were informed - is 1200 degrees C (just below the required temperature for finely knappable porcelain). Furthermore, a hollow inside the form is required for it to be printable, which is problematic for effective knapping activities. A separate company provided a sample of 3D-printed sandstone, which proved too thin and porous to be knappable (the current printing process is directed mainly at the printing of smaller decorative or novelty items such as figurines). Additionally, 3D printing is (still) quite costly, with a single printed blank (currently) costing >$100 USD. Despite the drawbacks with the currently available technology, 3D printing could still prove to be a useful process for the reliable production of standardized knapping blanks (perhaps even among the most reliable production techniques), assuming technical advancements are made and accessibility to the technology increases.

Beyond just 3D printing, ceramics generally provide one of the best options for achieving standardized knapping materials. Our observations already show that variation in the type of porcelain and the heating procedure can result in blanks of differing qualities, which more or less resemble different natural rock types (such as chert versus fine-grained basalt). More precise protocols - by using specific chemical formulations or heating procedures or combinations thereof - should be able to intentionally generate porcelains that match specific stones, thus increasing ecological validity of the porcelain used in experiments. Said validity can also be improved by increasing the heterogeneity of the blanks (e.g., by intentionally introducing controlled variation within the blank or adding in inclusions to mimic those found in rocks from natural outcrops). Besides increasing the external validity of the test materials (similar protocols could potentially be developed, e.g., with concrete), this would also create the potential of designing controlled experiments on knapping cognition where novice and expert knappers can be presented with the same exact raw material imperfections that they must then 'solve' (added a deeper level of control to pre-existing investigations into the varied strategies knappers employ to deal with raw material problems and their own knapping mistakes, e.g., Shelley, 1990; Torres & Preysler, 2020).

Conclusion

Controlled experimentation is an important tool for developing our understanding of the prehistoric knapping phenomenon, from the mechanics of fracture to the cognitive processes residing within the knapper (Johnson, 1978; Dibble & Whittaker, 1981; Toth, 1985; Schick & Toth, 1993, 1994; Dibble & Režek, 2009; Putt et al., 2014; T. Morgan et al., 2015; Moore & Perston, 2016; Lombao et al., 2017; Pargeter et al., 2019; Dogandžić et al., 2020; Snyder et al., 2022). By limiting the influence of confounds, we can more readily identify those processes involved under experimental conditions and, by extension, the processes guiding the formation of archaeological assemblages that we wish to better comprehend (Whittaker, 1987; Eren et al., 2016; Lin et al., 2018). One way that experimental control can be attained in archaeology is to produce and use knapping blanks that are standardized in material, size and/or shape (e.g., Sheets & Muto, 1972; Dibble & Whittaker, 1981; Dibble & Pelcin, 1995; Khreisheh et al., 2013; Režek et al., 2016; Speer, 2018; Dogandžić et al., 2020). Though not without its difficulties (we, the authors, attempted to secure a reasonable technique or source of blanks for months before settling upon a solution), said raw material standardization is a worthwhile pursuit. We have summarized all of the different raw materials and techniques we tested with the intent for use in our particular set of experiments; this summary serves as a basic framework for how standardized blanks can be selected and for how further techniques and materials might be imagined based on what we have already presented. Emergent technologies offer new possibilities for blank standardization, but these have not reached the degree of accessibility and refinement that would make them viable options in the present (e.g., 3D printing of glass and porcelain). There is no “one-size-fits-all” solution for blank standardization. No single material or technique or blank design can meet the demands of every thinkable research question in this field. Instead, researchers must balance the various facets involved in experimentation, from the research question to the accessibility of resources, the inherent costs of materials and production techniques, among others. It should be ensured that the materials and techniques used are both valid and reliable - to some extent. This guide is certainly intended to reflect this by providing information relevant to the planning of research projects, not only for new and prospective experimenters but also those with more extensive research. With more exploration of presently available options and future advancements potentially opening up new opportunities, we hope that future efforts will expand on what we have outlined here and that the expanding field of experimental archaeology will continue to improve as a discipline devoted to earnest scientific methodologies (Eren et al., 2016; Lin et al., 2018).

Acknowledgements

Access to a muffle furnace was provided by the Laboratory of Soil Science and Geoecology of the University of Tübingen, with specific help from Peter Kühn, Sabine Flaiz, and Rita Mögenburg. Thanks also go to Thomas Nieß and Karin Rein of the University of Tübingen Glass Blowing Workshop for access to machinery and advice on glass production. Special thanks for his contributions to our work on knapping experiments go to Shannon McPherron. We would like to thank the valuable input of our colleagues Patrick Schmidt, Li, Elisa Bandini, Alba Motes-Rodrigo, Vera Thomas, Katerina Harvati, and Nico Michiels. Special thanks go to Kathryn Ranhorn, Cory Stade, and Patrick Cuthbertson for their advice on porcelain blank-making, to Klaus Nickel and Silvia Amicone, and to Sam Lin for his observations related to the milling machine.

IV. Naïve knapper study: materials and methods

Explanation of candidate contribution to collaborative work

This chapter is based on collaborative work. The text contains (modified) excerpts from the main and supplementary text of the following publication:

Snyder, W.D., Reeves, J.S., & Tennie, C. (2022). Early knapping techniques do not necessitate cultural transmission. *Science Advances*, *8*(27), eabo2894. doi:10.1126/sciadv.abo2894

Conceptualization: C.T., W.D.S., J.S.R. Investigation: W.D.S. Visualization: W.D.S. Formal analysis: W.D.S., C.T. Writing (original draft): W.D.S. Writing (review and editing): W.D.S., C.T.

Some of the remaining content is from as-yet unpublished, in-preparation work:

Snyder, W.D., Reeves, J.S., & Tennie, C. (in prep). Untitled naïve human toolmaking and tool use paper.

Planning and conducting experiments with human volunteers

Ethics and data protection

In order to conduct experiments involving human volunteers, it is necessary to acquire approval in advance from the responsible regulatory bodies. In our case, we were required to submit an ethics application to the Ethics Committee for Psychological Research of the Faculty of Science, Eberhard-Karls University of Tübingen (hereafter, the Ethics Committee). The Ethics Committee was generally responsible for reviewing the research methodology, including safety procedures, and the data protection framework, as well as all relevant documents that were to be distributed to the participants in the study. Additionally, the European Research Council requested that we obtain approval from the local data protection authority, Zentrale Datenschutzstelle der baden-württembergischen Universitäten (ZENDAS; *EN*. Central Data Protection Office for Universities in Baden-Württemberg. Overlapping with the responsibilities of the Ethics Committee, ZENDAS reviewed the data management and protection strategies that we were to implement in the study.

The initial submission to the Ethics Committee (by W.D.S. and C.T., minus earlier communications and submission by C.T. alone) was on 13.08.2018. We received decisions (officially, 'Ethikvotum') usually approximately two to three months after the documentation was sent to the Ethics Committee. Though we had received ethical approval already on 20.11.2018 (i.e., after the second submission), changes to our protocol (such as the inclusion of glass as our knapping blanks and recommendations from ZENDAS) required amendments to the ethics proposal, which we submitted on 14.02.2019 and which were promptly rejected. Obtaining ethical approval required extensive work and involved a rather convoluted timeline, due to continued dissonance between the recommendations and requirements of the Ethics Committee and ZENDAS (e.g., in multiple cases, ethics reviewers questioned the validity of data protection protocols that ZENDAS specifically required us to implement in our study). ZENDAS also involved a more informal process, whereby advising on the data protection considerations for the study was provided over email by a representative of ZENDAS. W.D.S. also attended an in-person meeting with ZENDAS representatives at their headquarters in Stuttgart on 24.04.2019, in order to expedite final approval from their side.

Full approval from ZENDAS was received on 12.06.2019. Full approval ("ethisch unbedenklich", *EN*. "ethically harmless") from the Ethics Committee for the proposal titled

“Study on Human Stone Tool Making and Using” was received on 15.07.2019 (nearly an entire year after the initial submission and thus one year into the doctoral studies of the candidate.

The review process by ZENDAS had the following primary results as relates to the actual implementation of the study:

- The participant documents were edited to fit within the framework of the General Data Protection Regulation (GDPR) (EU) and the Landesdatenschutzgesetz Baden-Württemberg (LDSG) (EN. State Data Protection Law).
- Video recordings were required by ZENDAS to be pixilated to guarantee the anonymity of study participants. For this purpose, the video editing software Sensarea was applied, the usage of which was specifically approved by ZENDAS.
- ZENDAS determined that audio (i.e., voice recordings) can never be completely anonymized, even if with audio editing software. As a result, audio was not allowed to be recorded – ensured by us by physically blocking the microphone jack of the camcorders.

Study participants were given pseudonymized codenames, so that their real identities and personal information were never directly connected to study data. Access to the original, non-anonymized data was restricted to the authors of the ethics proposal (W.D.S. and C.T.).

A hygiene concept was also developed for testing after the advent of the COVID-19 global health crisis, which was approved by all relevant authorities and implemented as soon as 15.07.2020.

Safety procedures and the impact of the COVID-19 global health crisis

Safety procedures were designed and finalized based on the requirements of and consultation with authorities at the University of Tübingen.

During testing, all participants as well as the experimenter(s) were required to wear a pre-defined set of safety gear as follows (when applicable: gear type, brand, safety rating):

- Gloves (Galilee Royaltec CUT II), varying sizes, Category II
- Clear face shield (Clearways CV83P)
- Adjustable leather aprons (Babimax Retardant Welders Apron)
- Leather leg guards (AS Arbeitsschutz 272-302), welding standards

- Overshoe boots (Honeywell Over Shoes 971), varying sizes

Gloves and leather aprons protected individuals from cuts and percussive damage from shards of sharp materials. Category II gloves are a category above the minimal level for protection against cuts and scrapes, like those involved in knapping. Leg and foot protection were also included to prevent damage (via cutting) to the legs and/or clothing of participants, and to prevent small flakes and debitage from falling into participants' trouser legs or shoes. The clear face mask prevented sharp flakes and debitage from injuring eyes, nostrils, or mouth.

For tests performed after the advent of the COVID-19 pandemic (i.e., starting with participant P4), participants were additionally required to wear disposable gloves under the cut-resistant gloves (available in three sizes) and medical face masks (disposable face masks were provided upon request if participants did not bring their own or wished to change masks mid-test). The clear face shield was disinfected with antimicrobial disinfectant spray before and after use at each session. The room was also ventilated for at least 30 minutes before and throughout testing. Participants were required to fill out a questionnaire regarding their symptomatic status, in compliance with university regulation and legal standards. Individuals who were symptomatic or recently exposed could not be tested. Pregnant women could also not be tested due to an inability to meet the official recommendations for maintenance of safe distance (three meters). It was not possible to carry out the originally intended longitudinal testing strategy (10 sessions totaling 22 hours), because of COVID-19-related restrictions and dropouts caused by the pandemic. The two totally naïve individuals (P11 and P14) were invited to further testing sessions, but ultimately, neither of the two were able to comply with our request. Given that the totally naïve participants – at least, at the time – were the focal point of the study, plans for additional testing were abandoned. Similarly, a planned collaboration for a follow-up neuroimaging study (using functional near infrared spectroscopy or fNIRS) was cancelled due to the unavailability of participants.

Although it was determined to be highly unlikely that injuries would occur given the safety gear set we used, further proactive measures were taken to ensure the safety of participants. The main experimenter (W.D.S.) completed first aid training prior to the beginning of testing. The course (*DE*. EHB Erste Hilfe Grundschulung) took place on 8.12.2018, at the Johanniter Unfall Hilfe e.V. in Tübingen, Germany. A first aid kit (*DE*. KFZ-Verbandkasten), purchased for the purposes of this study was present at all testing sessions in case of any injuries to the participants.

Due to health and safety concerns, in cases where participants spontaneously engaged in potentially dangerous behavior (e.g., attempting to lift heavy objects like the granite block, throwing objects), experimenters discouraged repetition of the behavior by informing the participants that any repetition would result in the premature cessation of the testing session for their dangerous conduct (with compensation only for the completed testing time). There were no cases in which a session was ended prematurely because of dangerous conduct, but there was one case (P18) of a session ending early since the safety equipment – specifically the face shield – could not be sufficiently fixed to the participant's head and repeatedly fell to the ground, therefore preventing the safety protocol from being fulfilled.

Under the above guidelines, throwing was considered potentially dangerous conduct in most cases. Though throwing was a coded behavior due to its relevance as a technique for creating usable flakes (e.g., as spontaneously developed by the enculturated bonobo Kanzi; Toth et al., 1993), participants were discouraged from throwing after their first attempt, as it was deemed a health and safety concern. It was up to the experimenter's judgment whether a behavior qualified as "throwing" and therefore potentially dangerous, or "dropping" which was not considered as dangerous conduct. Generally, behaviors, even if then discouraged or disallowed, were still coded and recorded throughout the test.

Pilot experiments

In order to finalize the experimental design and methodology used in the study, we conducted a series of pilot experiments with 'non-naïve' participants (i.e., fellow members of the working group). These 'non-naïve' trial runs were then followed by three pilots with external volunteers (recruited the same way as outlined below in "Participants and materials").

Though the data from these pilot experiments – due to the application of conditions other than what was present in the thirty proper experimental sessions – were not included in the final data for the publication from this study (Snyder et al. 2022), these did lead to important changes to how the experiments were subsequently conducted:

- After the first 'naïve' pilot experiment, we made significant edits to the experimenter script (Appendix III), as there were clear issues with how participants understood when and which solution behaviors were subsequently disallowed and also which parts of the apparatus can and should be manipulated (e.g., Pilot Participant 1 tried to use the Impossible Flake as a ‘screwdriver’ to remove screws from the top of the apparatus boxes, which destroyed the Impossible Flake and could potentially have led to damage to non-replaceable apparatus parts).

- After the second ‘naïve’ pilot experiment, we decided to provide only one river cobble, instead of using the small-medium-large system that we used in the first two ‘naïve’ pilots (as cobble/’hammerstone’ selection would introduce an additional behavioral variable that was not intended for investigation, and thus would introduce noise into our dataset).
- We also introduced a new rope that was more static than the previous climbing rope that we used for the second ‘naïve’ pilot.
- The third ‘naïve’ pilot experiment was key to the development of the protocol for when participants used edges on a ‘core’ to cut the rope. In the initial conception of the study, the core-tool would have been immediately confiscated from the participant, but in the development of our methodology, we reconsidered this, in order to allow for further reduction of the core-tools (thus, allowing the appearance of the core types of interest and also more relevant behavioral data that might enhance our understanding of how Oldowan and Mode I tools were used).
- In the pilot experiments, the participants were given the option between opening two different apparatuses (thus, also including the ‘membrane box’ that was used in Bandini et al., 2021c and Motes-Rodrigo et al., 2022). Pilot Participant 3 was able to reach into the opening of the membrane box and simply peel away the membrane to retrieve the award. We consequently decided to remove the membrane box completely from the finalized experimental design for the study.
- Pilot Participant 3 was the only one of the three to receive the questionnaire on past experience (Appendix III). At this stage, we decided to provide all participants the questionnaire, regardless of whether they produced and used a cutting tool or not.

Logistical considerations

Besides the extensive work involved in acquiring approval from the responsible ethics and/or data protection authorities, there are a great number of other logistical components that need to be arranged and handled prior to and during experimental research with human volunteers.

Among these is the selection of suitable (i.e., valid and reliable) raw materials for use in the knapping experiments (see “Chapter III: An overview of standardizable raw materials for controlled human knapper experiments”). After much trial-and-error in the selection of the raw materials, there was then the task of obtaining sufficient raw materials for the full duration of the study. Indeed, it took several months for the complete delivery of the glass hemispheres from the available online glass wholesalers.

Our research group did not have exclusive access to a testing space and, thus, needed to make arrangements to use a set of rooms in the collections of the Zooarchaeology Working Group (A.B-L. and B.S., acknowledgements). Importantly, we needed to identify a space that did not in some way contain direct hints to the purpose of the experiment (e.g., posters with images of stone tools). Before each test, the experimenter ensured that this remained the case throughout the experiment, thereby controlling for exposure to the pertinent know-how and any related stimuli that could influence the participants’ naivety unnecessarily.

Scheduling with the participants had to account for the usage of the space by other members of the institute. This additionally meant that the testing setup needed to be unpacked and packed on each day that experimental sessions took place. Other than the four hours of testing time for each session (up to five hours in total when including the arrival of the participant and introduction to the testing paradigm), at least one hour was required for the combination of setting up and cleaning for each session. For days in which there were two consecutive testing sessions, the day’s total time investment (excluding follow-up tasks such as document digitization and anonymization) could amount to ten or more hours of test-related tasks. There was also the occurrence of sunken costs for setting up and cleaning up the experiment due to no-shows (i.e., individuals who scheduled appointments to participate but then – without prior indication – did not arrive at the testing location) or abandoned tests. Abandoned tests occurred, e.g., when a pregnant woman wanted to participate but it was barred due to the COVID-19 hygiene concept or when, in the case of P18, the face shield could not be appropriately affixed without repeatedly falling off. One session was completed, but the participant (P20) did not submit written consent (despite having verbally indicated to the experimenter that the written consent had been completed and sent by email). As a result, the data that was collected could not be used, and all video recordings and personal/demographic information for the participant was deleted and/or shred.

Participants

Thirty participants were recruited via online and newspaper classified listings in a town in southwestern Germany. Participants were requested to fall within the age range specific by the advertisements (18 to 55 years of age). Due to data protection concerns, we did not ask participants to directly report their age, instead only asking for the decade of birth in the pre-study demographic questionnaire. Participation in the study was contingent upon good eyesight or good corrected eyesight, right-handedness, unrestricted physical ability (i.e., the freedom to move arms/hands and strength to perform the task), no diagnoses for neurological or psychiatric illnesses, and no active psychoactive drug use; participants were asked to self-filter according to these criteria prior to testing. Two participants were tested but their data were excluded (due to safety concerns and experimenter error, respectively). A final sample of twenty-eight tested participants (fourteen female and fourteen male) are reported here.

Participants in this study were "WEIRD" (referring to people from societies that are "Western, Educated, Industrialized, Rich, and Democratic"; Henrich et al., 2010). This choice provides commensurability with previous knapping studies, which likewise used participants from WEIRD populations. WEIRD individuals are nonetheless appropriate for this type of study (i.e., an Island Test), since they are generally unlikely to be exposed to knapping experiences in their daily lives.

Materials

At the start of testing, each participant was provided with one spray-painted glass hemisphere (10 cm in diameter and 4 cm in height; note that hemispherical blank forms or "split cobbles" are known from the Oldowan: Toth, 1985), one locally sourced river cobble, and a large rectangular granite block (Figures 16 and 17, Table 6). Participants could receive replacements for the hemispheres and cobbles if the materials were deemed exhausted according to the participant's own evaluation. The puzzle box afforded the use of tools and consisted of a box with an enclosed reward accessible by severing a rope - a 'tendon' box (Figures 16 to 20; Bandini et al., 2021c; Motes-Rodrigo et al. 2022). Due to the length of testing, both participant and experimenter(s) were given chairs. All tests were recorded using three separate digital camcorders, mounted on tripods at designated spots around the testing area. Participants were remunerated with €12 per hour plus any attained reward money.

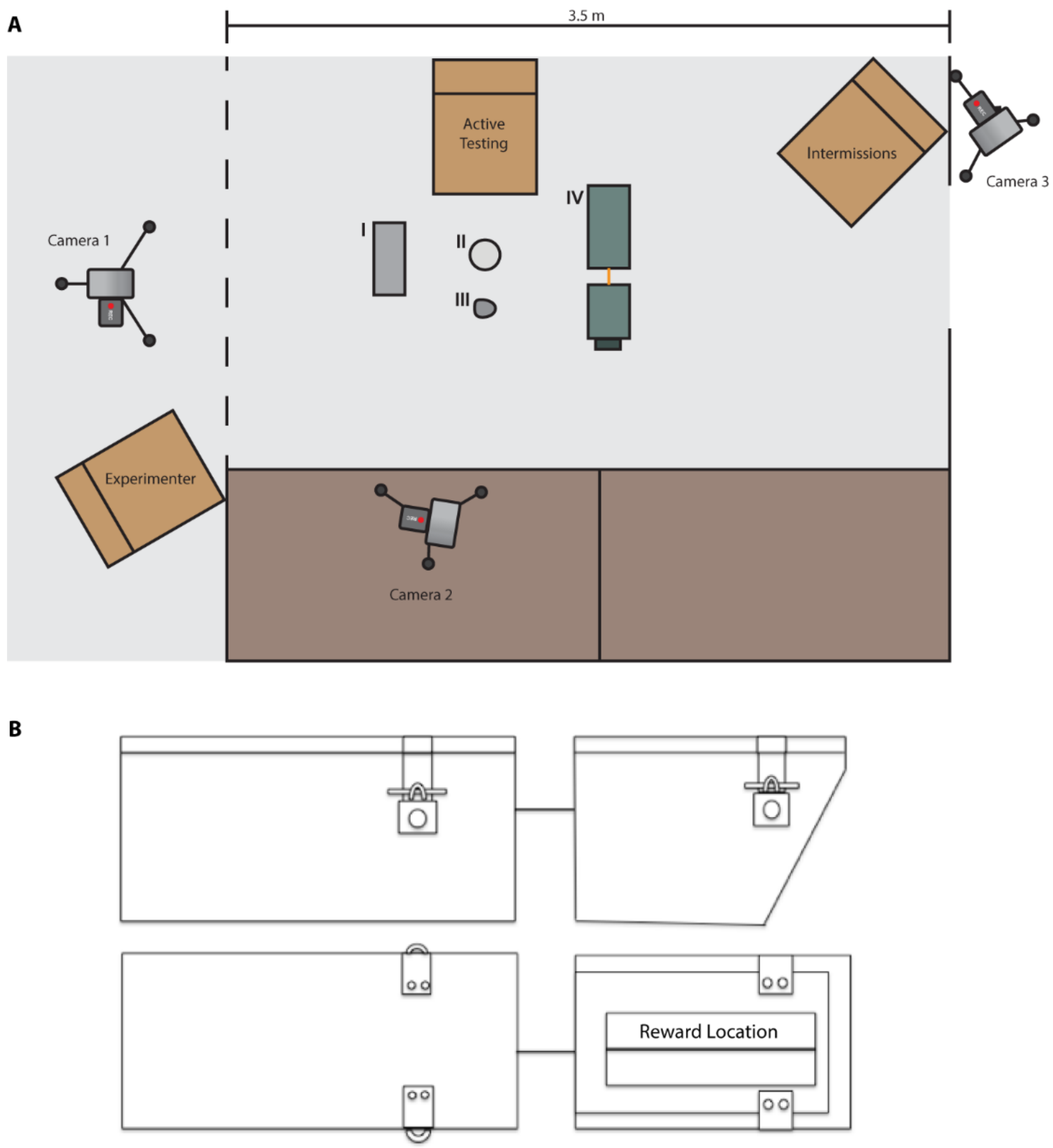


Figure 16 The experimental apparatus. (A) Basic layout of the experiment, including the main testing materials like the (I) granite block, (II) glass hemisphere, (III) river cobble, and (IV) puzzle box. (B) Schematic of the puzzle box with side (top) and top (bottom) views.

Table 6 Weights for hammerstones from Olduvai Gorge Oldowan sites (de la Torre & Mora, 2005) and river cobbles from this study (Snyder et al., 2022).

Site Name	Mean mass (g)	Std. deviation
DK	462.90	203.911
FLK Zinj	351.36	160.68
FLK North 1-2	390.56	151.267
Experimental selection	378.24	102.04

Glass hemispheres

Clear, uncolored glass hemisphere paperweights were purchased from online wholesalers. The hemisphere was standardized in terms of shape and mass (diameter: 10 cm, height: 4cm, mass: 600g). The glass blanks were spray-painted with light grey spray-paint to create an artificial ‘cortex’ for purposes of concealing the material properties from the test participants and for improvement of attribute measurements post-test.

Overall, thirty-five glass blanks were provisioned to participants.

River cobbles

Participants were provided a river cobble at the start of the test. River cobbles were purchased from a local construction supplies firm (Natursteinpark Tübingen). These were then selected for use as hammerstones (at least, as predicted by the experimenters) in accordance with the means of masses reported for Oldowan sites at Olduvai Gorge Beds 1 and 2 (Table 6; de la Torre & Mora, 2005), since the data were published and already used to design hammers used for knapping experiments with great apes (Bandini et al., 2021c; Motes-Rodrigo et al., 2022). Standard deviations were compared as well to make sure that the experimental selection was similar in variation to the archaeological hammerstones.

Overall, thirty-eight cobbles were provisioned to participants.

Granite block

A large rectangular granite block (34cm x 15 cm x 13 cm; 16.7 kg) was provided to the participants. The granite block could be used by the participants as, for example, an anvil in bipolar technique or a passive hammer in the eponymous technique (de la Torre, 2019; Harmand et al., 2015; Putt, 2015; Toth & Schick, 2011, 2018; Wynn & McGrew, 1989).

Methods

Experimental design

Participants were not given any relevant information (verbal, written, or visual) about knapping or stone artefacts before or during the experiment. The participants were merely instructed to procure the reward from the puzzle box, using the available materials in the testing space. Participants were tested individually to further prevent any transmission of information and to ensure independent replication. The experimenter abided by a pre-

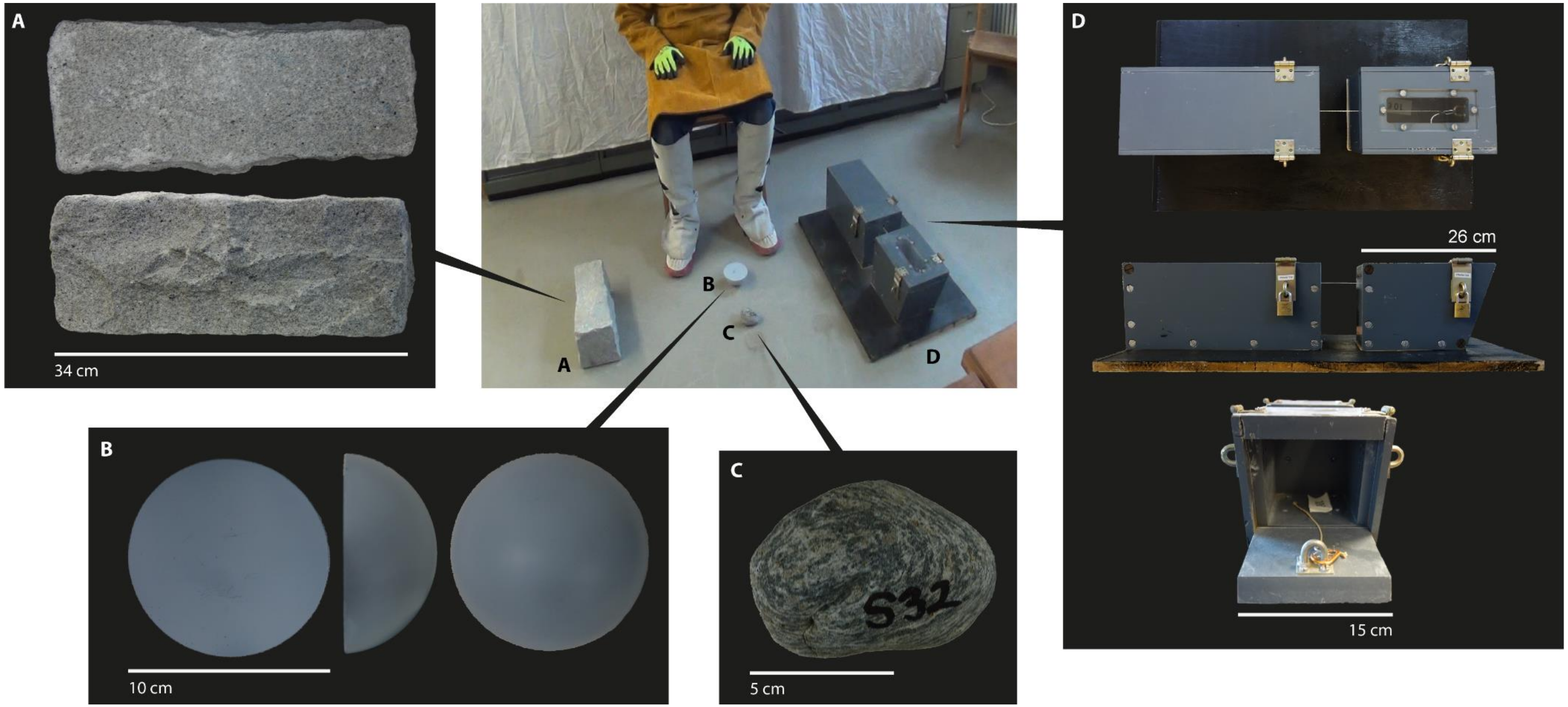


Figure 17 The experimental apparatus at testing start (here, for Participant 3; seated in the chair at the center, top), including (**A**) the granite block, (**B**), the painted glass hemisphere, (**C**), the river cobble, and (**D**) the puzzle box. This is the view as seen from Camera 2 (Figure 16). Photo Credits: William D. Snyder and Jonathan S. Reeves, University of Tübingen

determined script to avoid phrasing that might reveal the goals of the experiment (e.g., the experimenter referred to the cores, flakes, and other materials exclusively as “objects” and never specifically mentioned “cutting”; see Appendix III). Overall, this setup allowed us to test whether participants would produce and use glass or stone cutting tools by any variant of knapping technique, without compromising their naivety. This also ensured that the spontaneity of their technique innovations could be verified.

Procedure

Participation was contingent upon informed consent and fulfilment of eligibility criteria (ethical approval by the Ethics Committee for Psychological Research, University of Tübingen, confirmed on 15.07.2019). Pre-study information for the participants only indicated that this would be a study on “human problem-solving abilities”, and did not mention stone tools, their form or manufacture techniques. This deception was necessary because transparency about the study’s aims would have compromised the participants’ naivety.

Due to health and safety regulations, participants and experimenter(s) wore safety gear and were not permitted to remove any of it while in the testing area. Each participant was tested for a maximum four-hour test (two tests ended early by participant request). Just prior to testing, participants were given explanations of the testing setup, including safety precautions and general study rules (e.g., no access to smartphones, no questions about specific solutions). The experimenter also clarified the mechanics of the puzzle box (i.e., that a visible rope prevented a door from being opened, thus further blocking access to the reward inside; this was communicated, however, using only gesture and vague language, see Appendix I and III). After the initial study explanation, the experimenter signaled to start, at which point the participant could begin pursuing solutions. If participants used a solution other than creating and using a cutting tool, the solution was noted and named live by the experimenter, and the participant was then instructed not to repeat this solution and to instead pursue other (unnamed) avenues of opening the puzzle box. The puzzle box was then reset and rebaited. The puzzle box was also reset and rebaited upon successful creation and usage of a cutting tool (either as a detached piece or cutting edge on a core).

Used detached pieces were retrieved by the experimenter and bagged and labelled, while successfully used edges on a core were visibly marked by the experimenter (with a pen) and

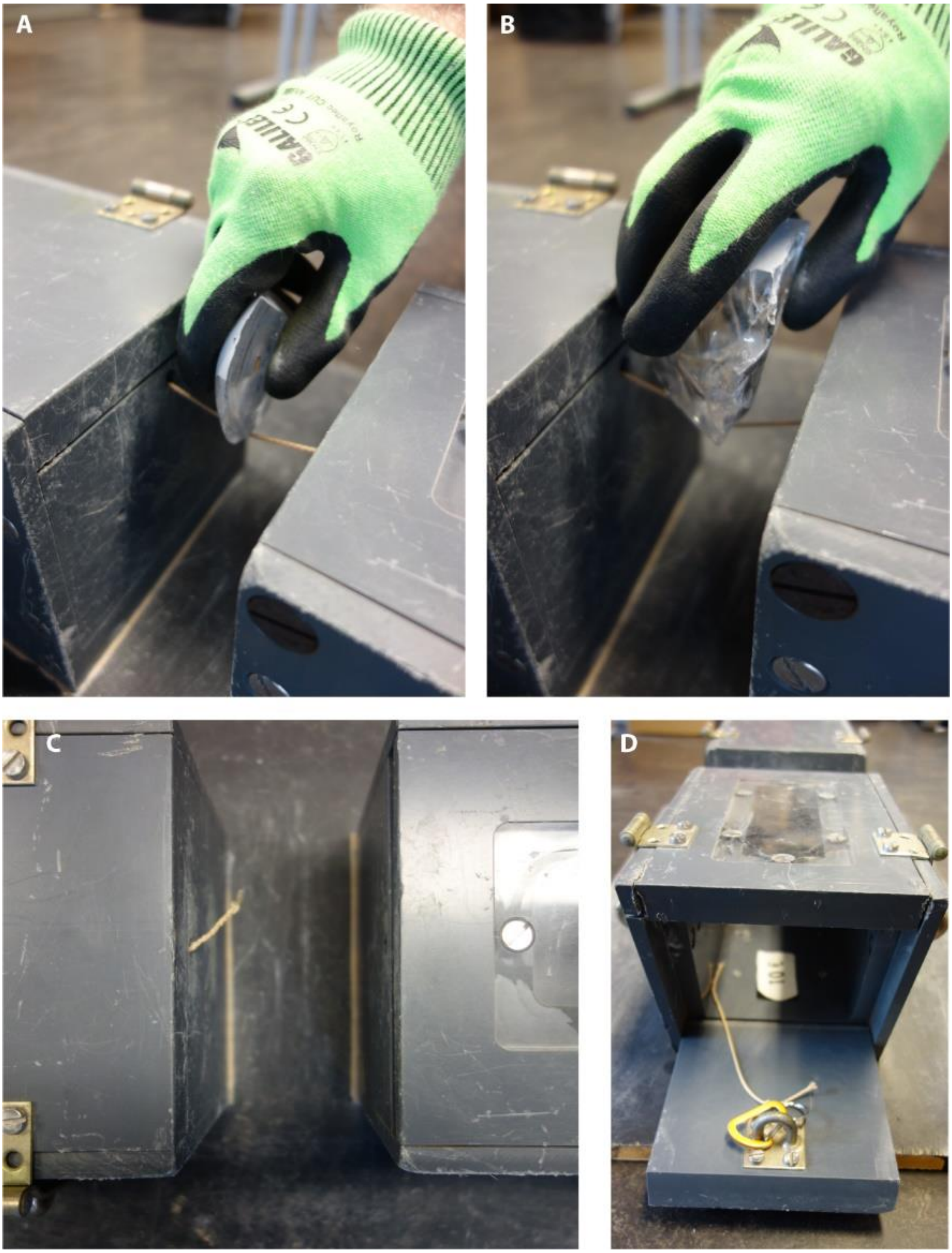


Figure 18 Cutting tool use. (**A**) The use of a glass cutting tool (flake) to sever the rope and thereby open the puzzle box. (**B**) The use of a glass cutting tool (here, a core-tool) to sever the rope and open the puzzle box. (**C** to **D**) Opening of the puzzle box as a consequence of the tool use action. (**C**) Breakage of the puzzle box rope. (**D**) Opening of the freed door of the front compartment, allowing access to the reward token. pictures are derived from re-enacting by the experimenter, in order to demonstrate the use of a tool and the mechanism of the puzzle box with clarity that could not be achieved from the captured video stills of the actual tests. Photo Credits: Claudio Tennie and William D. Snyder, University of Tübingen.

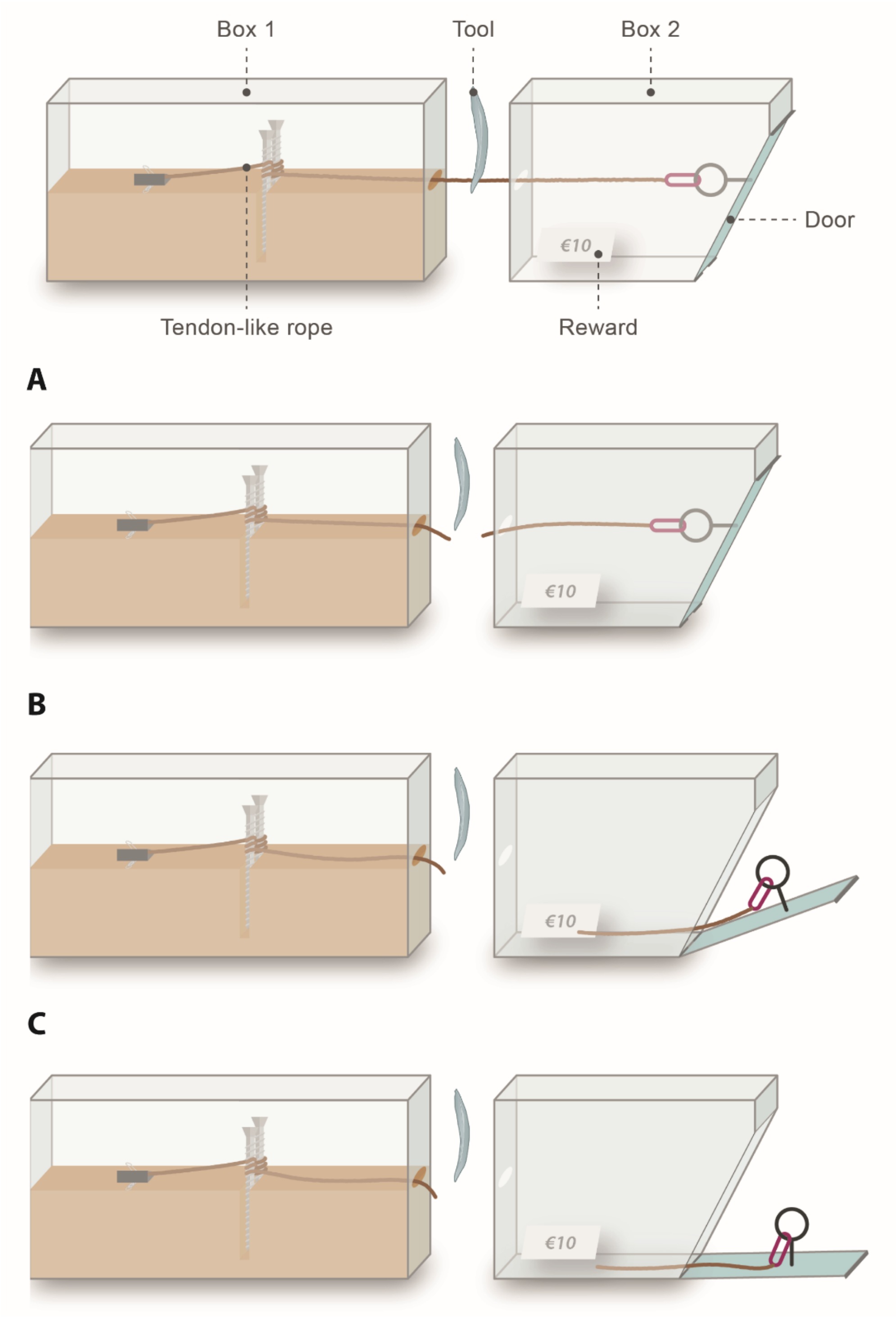


Figure 19 Mechanism for opening the puzzle box apparatus. The apparatus consists of two main compartments (i.e., boxes). A tendon-like rope is led from Box 1 through into Box 2 and then used to fasten shut the door at the front of Box 2. As such, to open the door and access the hidden reward (which is visible to the participant via a window at the top of Box 2) (Fig. S1), the participant must somehow break or sever the rope. In this example, a flake (here, not to scale) was produced from the glass hemisphere and used to cut the rope (**A**), thus relieving the tension in the rope and opening the door of Box 2 (**B**). With the door now open, the participant can reach in and remove the reward slip (**C**). In other cases, participants used the core as a cutting tool or opened the door of the puzzle box using a non-cutting solution (Fig. S2 and supplementary text). Graphic created by N.M.M-G., see acknowledgments.

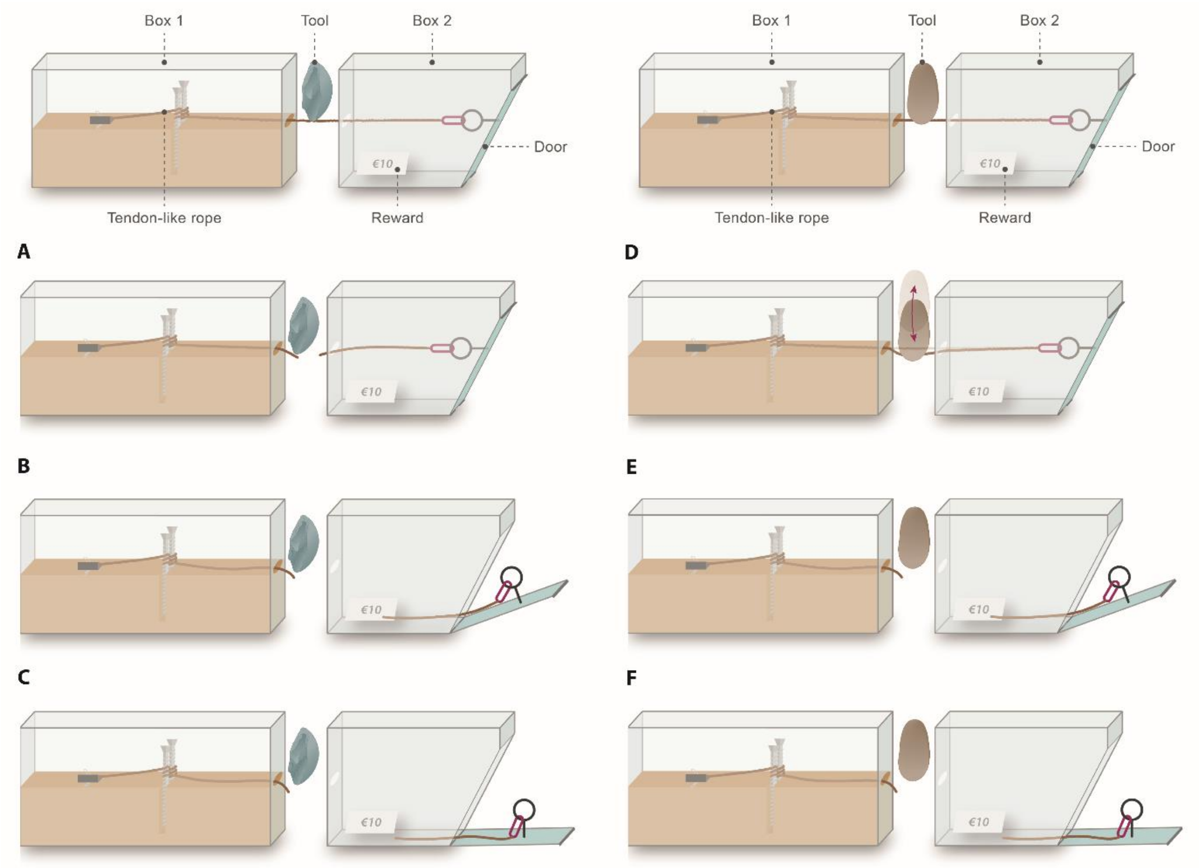


Figure 20 Alternative pathways to solving the puzzle box. Cutting with a flake-tool (Figure 19) is not the only potential solution for the puzzle box. Here are two examples of said alternative solutions. In the first example (**A** to **C**), cutting edge created on the core (i.e., a core-tool) is used to cut the rope and therewith make the reward accessible by opening the puzzle box. In the second example (**D** to **F**), a non-cutting solution is applied. In this case, the river cobble is used to hammer on the rope with a percussive action until the rope breaks, thus opening the door and making the reward accessible to the participant. Some other examples of non-cutting solutions that participants applied included rubbing with a blunt edge (e.g., the river cobble or the round face of the glass hemisphere), picking apart the rope threads using their fingers, and stepping on top of the rope. Graphic created by N.M.M-G., see acknowledgments.

verbally designated as unusable. This procedure simulated edge wear (as occurs in naturalistic conditions) and thus encouraged the production of new edges. If the participant did not produce a cutting tool after two hours, they were provided an 'impossible flake' (i.e., a tool with a functional shape – a wedge-like triangular prism made of glass and spray-pained – but impossible to recreate via knapping) for exactly one trial. Puzzle boxes were baited with paper slips designating monetary rewards (Figures 16 to 20). Participants received this monetary reward (as credit) for each time that they opened the puzzle box, regardless of the solution type.

After approximately five minutes of fruitless toolmaking attempts and/or total inaction (at any point during testing), participants were asked if they would like to receive a new blank. If participants requested a new blank (i.e., deemed exhaustion of the previous one), the old core and debitage were all collected and bagged with labels. Otherwise, collection, bagging and labelling of cores and debitage occurred at the end of the test.

In order to ensure that participants were naïve and remained naïve before and during testing, it was necessary to collect data on the participants' previous experience with stone tools after the test. Thus, once the testing time was complete, participants filled out a questionnaire about their pre-study experiences – of any type – with stone tools (a 'naivety check'; DE. Fragebogen zu den Vorerfahrungen der Studienteilnehmer, EN. Questionnaire on the Past Experiences of Study Participants, see Appendix III).

Participants who, during their entire test, did not make their own cutting tool at all or who made cutting tool(s) but did not use them were given a questionnaire aiming to determine the cause of their lack of success (DE. Post-Studie Fragebogen (von Teilnehmern, die nicht erfolgreich waren), EN. Post-Study Questionnaire (for participants who were not successful)). This questionnaire was in two parts: an initial section and a follow-up immediately after the participants were provided information on the goal of the task and stone tools (see Appendix III).

A third questionnaire was given specifically to participants who, during their entire test, created at least one cutting edge but failed to utilize these to open the tendon box (DE. Weitere Fragen im Falle von einer Teillösung der Aufgabe, EN. Additional questions in case of a partial solution, see Appendix III).

Data

Behaviors were coded live by the experimenter using a paper coding sheet. After the test ended, video recordings were processed (angle-selection and anonymization) and then coded. First, behavioral bouts (actions or action sequences with a definable start and endpoint) were identified as involving toolmaking, tool use (or other solutions to the apparatus), or simultaneous toolmaking and tool use. We coded events as potential toolmaking, both when actions did lead or could have led to fracture of an object and, with it, the creation of a potential cutting tool. Confirmed toolmaking events, on the other hand, were only those where the fracture of an object could be determined visually from the video recordings (Figures 16 and 24). The dichotomy of potential versus confirmed toolmaking was implemented as the intentions of the participants are not strictly known (unlike other studies, where the intentions are prescribed by the experimenters and therefore reasonably identifiable; Cataldo et al., 2018; Lombao et al., 2017; T. Morgan et al., 2015; Pargeter et al., 2021; Stout et al., 2019). Toolmaking bouts were further assigned to one of four early knapping technique categories: freehand, passive hammer, bipolar, and projectile. Freehand, passive hammer, and bipolar techniques are all specifically associated with Oldowan technology (Toth & Schick, 2011, 2018), while passive hammer (Harmand et al., 2015) and bipolar technique (Putt, 2015; Wynn & McGrew, 1989) are also associated with purported and hypothetical pre-Oldowan technologies. Projectile technique is not typically associated with the Oldowan proper but has been cited as a potential pre-Oldowan behavior due to its expression by the captive bonobo Kanzi and may have played a role in the innovation of knapping by naïve Oldowan hominins (Putt, 2015; Toth et al., 1993). We used a "constellation-based" approach for the assignment of toolmaking bouts to one of the technique categories; this method was based on earlier approaches (de la Torre & Mora, 2005) but modified further so that technique classification was fully neutral in terms of intention, outcome, and mechanic (e.g., percussion, friction, pressure). In this approach, objects are identified using neutral terms like active, passive, auxiliary and target element, instead of already interpretative archaeological terms (e.g., 'hammer' or 'anvil').

- In freehand knapping, the agent (here, study participant) uses an object held in one hand (an active element) to induce a possible transformation in another object that is held in the other hand and/or supported against the body of the agent (a passive element or a second active element).
- Passive hammer technique involves one object held by the agent (an active element) being moved (mainly) to be transformed by another (passive element) that is either stationary on the floor or even the floor itself (note that passive elements can also be transformed in this scenario).
- Bipolar technique involves one object (active element) held and moved by the agent to induce a potential transformation in another (passive element), which is stationary and resting on/supported by another object (auxiliary element).
- Projectile technique involves one object (active element) being dropped onto (i.e., leaving body contact) or thrown toward another object or the floor (passive element).
- Simultaneous toolmaking and tool use can involve a similar constellation of elements to the previous categories (in this study, only passive hammer technique was not a technique for TMU), but also the target element, meaning it is the simultaneous production of a tool or cutting edge and use of a tool.

Toolmaking mechanics were coded as being percussion, friction, or pressure, while outcomes were coded in terms of whether there was fracture of an object and which object was fractured in the bout.

If a participant used a technique in ≥ 50% of toolmaking events, this was defined as a preference.

For possible solution events, the bout was coded as either successful (the puzzle box was opened during the bout) or unsuccessful (the puzzle box was not opened during the bout). All possible or attempted solutions were coded as having a solving technique. Bouts were coded as CUT if there was any sort of cutting, slicing, or chopping of the puzzle box rope, either using an edge created by the participant (i.e., a core-edge or a detached piece/flake) or with the impossible flake. RUB-S was any rubbing or friction-based action on the puzzle box rope with a pre-existing acute-angled edge (e.g., an unworked edge of the painted glass core), while RUB-B was any rubbing or friction-based action with a rounded or blunt edge of an object. BOD was any solution attempt without the assistance of an object, i.e., wherein the participant tried to break the rope of the puzzle box with their hands or with their foot. PRS was any solution attempt action whereby an object was pressed against or down on the rope, pushed against the rope, or placed just under the rope with a corresponding pulling up action in order to weaken and/or break the rope. PER was any percussive action direct onto the rope with at least one object or the blunt ends/surfaces of two objects. SYN refers to the specific circumstance during bouts coded as simultaneous toolmaking and use, where the fracture of an object occurred but it is unclear whether the rope was broken, e.g., by cutting/slicing of the new cutting edge or through the percussive forces between the two objects. Finally, OTH was the category used for any actions that were identifiable (as opposed to unidentifiable actions, i.e., UNK), but which did not fall into the preceding categories. All objects involved in solution attempts or tool use bouts were also coded (as tool1, tool2, and tool3, ranked from most to least active or vertically from top to bottom).

The proportion of total cutting attempts with a debitage element out of the total cutting attempts was calculated for all participants. If the proportion was greater than 0.6, the participant was identified as having a preference for using ‘flake-tools’. If the proportion was less than 0.4, the participant was identified as having a preference for using ‘core-tools’. If the proportion was between 0.4 and 0.6, the participant was identified as being an ‘opportunistic’ cutting tool user.

For interobserver reliability calculation, a subset of data (n = 7 participants; 25.0% of participants) was coded by a hypothesis-unaware individual (J.K.) according to the above criteria. The coding from the experimenter and the reliability coder were then compared.

Participants' post-test questionnaire answers were used to rank the naivety of the participants, with a minimum rank of 0 for totally naïve individuals (with no theoretical or practical knowledge of stone tools), a rank of 1 for individuals with basic conceptual knowledge of stone tools (i.e., conscious of the existence of stone tools), a rank of 2 for individuals who had seen a stone tool or a depiction of a stone tool, a rank of 3 for individuals who had seen the production of a stone tool, and a maximum rank of 4 for individuals with at least some previous hands-on experience with stone tools and knapping. Participants with naivety ranks between 0 and 2 were considered technique-naïve (i.e., naïve to the specific know-how of early knapping techniques), while naivety ranks of 3 and 4 qualified participants as non-technique-naïve.

All detached pieces used by the participants as tools were immediately collected, labelled, and bagged for later analysis. Cores and all associated debitage (no matter how small) were collected and labelled either at the end of the test or if the core was exhausted. Used tools of any size and debitage elements equal to or above 2 cm in maximum dimension (Key & Lycett, 2014) were subject to categorization as flakes or angular fragments (Braun et al., 2019) and an attribute analysis.

Artefact analysis included the following variables: artifact type (flake, angular fragment, glass core, or cobble), mass (g), maximum dimension (mm), width (mm), and thickness (mm). For flakes, platform depth (mm) and exterior platform angle (°) were also measured. Maximum dimension, width, thickness, and platform depth (*PD*) were measured using digital calipers. For exterior platform angles (*EPA*), three separate measurements of the angle were made with a goniometer and then averaged (all reported EPAs are thus averaged values). Values for PD and EPA from the experimental flakes (separated into naivety groups) were compared with a reference dataset (Režek et al., 2018) from Oldowan sites (for our purposes, flakes that were labelled in this dataset as "Oldowan", "Developed Oldowan", and "Karari" were combined into one Oldowan technological category). Further comparisons were made between the full experimental sample of flakes with the same reference dataset.

For the cores, basic technological analysis was performed to elucidate the reduction strategies implemented by the participants (de la Torre, 2011) and the basic typology of the final cores

(Leakey, 1971; Toth, 1985). Cores were matched to the classical Oldowan types to which they most closely resembled based solely on how the types are defined morphologically and technologically (i.e., without any assumptions about intention). Additional variables relevant for technological analysis of cores included flake scar counts, step/hinge fracture counts, and battering (visual indications, yes or no). The post-test technological analyses were then compared to the technological actions indicated by the live and post-test behavior coding.

For further characterization of the technology produced by the naïve knappers and comparison with assemblages from the Early Stone Age of East Africa, we also conducted principal component analysis using the variables and data first applied by Braun et al. (2019). Metadata for East African Early Stone Age sites was compiled by Braun et al. (2019; Delagnes & Roche, 2005; de la Torre, 2004; de la Torre & Mora, 2013; de la Torre et al., 2003; De Lumley & Beyene, 2004; Gallotti & Mussi, 2015; Harmand et al., 2015; Hovers, 2009; Ludwig, 1999; Mora & de la Torre, 2005; Proffitt et al., 2016; Stout et al., 2010). The principal component analysis included the following variable: percentage of unifacial cores in total core assemblage, percentage of bipolar cores in total core assemblage (here, we used the percentage of toolmaking events that involved bipolar technique), percentage of cores in the complete assemblage, percentage of angular fragments out of overall detached pieces, average core maximum dimension, ratio of average flake size to average core size, average flake scar count per core, ratio of flake scar count to log (base 10) of mean core size, average flake maximum dimension, average flake thickness, and percentage of percussive tools (objects in entire assemblage that show signs of percussive activities). For the final variable, we did not classify objects with battering or crushing only at the platform edges as percussive objects. For missing values, a multivariate normal imputation (via least squares prediction and with shrinkage) was applied.

In order to evaluate the effects and outcomes of different knapping techniques, we adapted the analyses for productivity, expediency, and efficiency employed by Putt (2015) in their study of knapping ‘precursor behaviors’. We follow also the definitions used therein for the three categories of toolmaking measurements: productivity “measures the output of usable, functional tools”, expediency “measures the ease and time it takes for productivity to occur”, and efficiency “measures the relationship between the effectiveness of raw material exploitation and conservation of energy” (Putt, 2015, p. 54). For productivity, we considered three different measurements: total flanks produced per blank, total flakes produced per blank mass, and the ratio of flake maximum dimension and flake mass. For expediency, we considered four different measurements: total flakes produced per second (here, per second of total duration of toolmaking bouts), total debitage mass produced per second, total cutting successes per second, and total cutting successes per toolmaking bout. Finally, for efficiency, eight measurements were applied: percentage of blank mass converted into flakes, percentage of blank mass converted into unused debitage (i.e., detached elements), percentage of blank mass converted into unexploited core, proportion of flakes among debitage elements, total count of successful debitage per total debitage, total cutting successes per blank mass, total flakes per toolmaking bout, and total successful debitage per toolmaking bout. For these measurements and analyses, we did not identify any detached pieces from river cobbles as flakes, to control for raw material and because of inconsistency between observers in identifying conchoidal fracture from river cobble detached elements. Due to the specific research questions and testing conditions used in this study, we were often unable to maintain commensurability and therefore did not use all of the same measures for productivity, expediency, and efficiency used in Putt (2015). For example, the expediency measure of total time duration to completely exploit a core was not included, because many of the cores were not completely exhausted (i.e., the testing session ended before the eventuality of our core exhaustion protocol came into play).

Statistical analyses

Results were considered statistically significant at an α level of 0.05.

For the results initially reported in Snyder et al. (2022), statistical analyses and graph-building were performed using RStudio (RStudio Team, 2020). Due to the non-normality of the data and the small size of the sample for the different naivety categories, non-parametric tests (Kruskal-Wallis test for one-way analysis of variance and Wilcoxon rank-sum tests for pairwise comparisons) were used. Tests were two-sided (per the default setting in RStudio).

Cohen's kappa tests for interobserver reliability were used to measure agreement between the coding data of the experimenter (W.D.S.) with the coding data of the naïve secondary observer (J.K., acknowledgments). Cohen's kappa tests were used on count data (frequency of a behavior or behavioral character) and Boolean data (presence/absence of a behavior or behavioral character). Similarly, a Cohen's kappa test was performed for interobserver reliability of flake identification, involving a comparison between the typing data from the experimenter (W.D.S.) and the hypothesis-naïve secondary observer (A.F., acknowledgments).

For the extended results reported first in this dissertation, statistical analyses were performed in JMP 16.0.0 (JMP, 2021). Graph-building was performed in Past 4.11 (Hammer et al., 2001). For comparison of flake masses between archaeological and experimental flakes, a non-parametric test (Wilcoxon rank-sum test) was used, because the distributions for both samples were non-normal. Statistical significance for the remaining flake variables was determined using one-way Student's t-tests. Statistical significance of productivity, expediency, and efficiency measures was determined with Kruskal-Wallis tests for one-way analysis of variance, because of uneven sample sizes, unequal variances, and non-normal data distributions. Follow-up each pair comparisons were conducted using the Wilcoxon rank-sum method. For tool use preferences, most variables were analyzed using ANOVA analyses for variance, with pairwise comparisons explored using Student's t-tests. The exceptions were flake mass and platform depth, which were non-normally distributed; scar counts on cores, which were non-normally distributed and had insufficient sample sizes; and flakes produced per blank, which had insufficient sample sizes. These variables were analyzed with Kruskal-Wallis tests for one-way analysis of variance, with pairwise comparisons analyzed with Wilcoxon rank-sum tests.

V. Naïve knapper study: results

Explanation of candidate contribution to collaborative work

This chapter is based on collaborative work. The text contains (modified) excerpts from the main and supplementary text of the following publication:

Snyder, W.D., Reeves, J.S., & Tennie, C. (2022). Early knapping techniques do not necessitate cultural transmission. *Science Advances, 8*(27), eabo2894. doi:10.1126/sciadv.abo2894

Conceptualization: C.T., W.D.S., J.S.R. Investigation: W.D.S., Visualization: W.D.S., Formal analysis: W.D.S., C.T., Writing (original draft): W.D.S., Writing (review and editing): W.D.S., C.T.

Some of the remaining content is from as-yet unpublished, in-preparation work:

Snyder, W.D., Reeves, J.S., & Tennie, C. (in prep). Untitled naïve human toolmaking and tool use paper.

Participant naivety

As indicated by the posttest questionnaires, 25 of 28 participants were technique naive (i.e., had no knowledge of or previous experience with any knapping technique; naivety levels 0 to 2; Table 8). Of these 25 technique-naive participants, 22 participants had only seen stone tools before (naivety level 2); one participant had heard of stone tools (level 1); and two participants (participants 11 and 14, henceforth P11 and P14) had been totally naive concerning stone tools, stone tool types, and knapping techniques (level 0) before testing.

Several participants who had admitted to seeing stone tools before provided drawings approximating what they had previously experienced (see Figure 21). The depictions were of objects or artifacts that do not superficially resemble Oldowan or Mode 1 artifacts. These include a number of hafted tools (identified by participants as, e.g., hammers and axes), millstones, and arrowheads, as well as two that might be identified potentially as hand-axes or Achealean-type bifaces (though the participants name them only as ‘stones’ and do not use the aforementioned archaeological terminology).

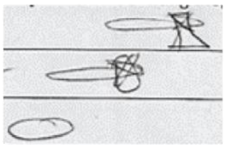
Over the course of testing, several participants also vocally expressed surprise during or just after their first bouts of successful flaking (e.g., P7, P10, P19, P26). Additional participants made other comments and conjectures that were noted in the live coding, due to particular relevance to their naivety and understanding of what was taking place in the experiments (see Table 7).

Main results (Snyder et al., 2022)

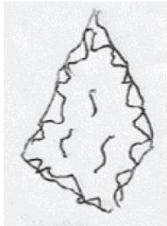
Despite the naive nature of 25 participants regarding knapping techniques, all 25 of these spontaneously used knapping techniques, 22 of whom spontaneously produced and then used cutting tools (two others made potential cutting tools, and one of those two used cutting tools after having been given the impossible flake, i.e., non-spontaneously; Table 9). Both totally naive (level 0) participants individually developed early knapping techniques (Figure 22) and used the resulting artifacts as cutting tools (Figures 22 and 24 to 26).

Overall, we recorded 1580 potential toolmaking events (i.e., where the actions therein could have or did lead to the fracture of an object). Of these potential toolmaking

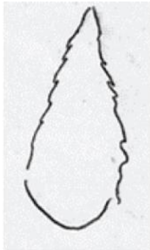
P1
"Dark stone, which was worked on all edges"



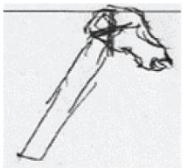
P2
"Axe, hammer, stone"



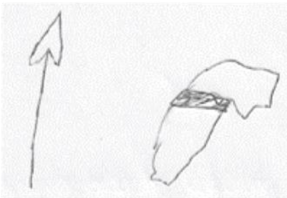
P8
"Arrowhead or type of knife"



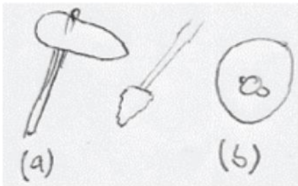
P17
"wedge-shaped stones, not sure if one side smooth and one wavy or both sides smooth; hand-sized"



P22
"Hammer or axe"



P26
"Stones, which were bound with leather straps or sharpened stones, which looked like arrows""



P27
"Axes/hatchets made of stone (a) / millstones (b)"

P30
"Axe, hammer?"

Figure 21 Drawings of stone tools submitted by study participants as part of their responses to the questionnaire on past experiences.

Table 7 Selected utterances and remarks of participants noted in the live coding sheets.

ID	Date	Time	Live coding quote or note
P1	28/1/20	11:29	"I don't want to break glass."
P7	14/8/20	14:42	Expressed shock due to seeming unintentional flaking
P10	19/8/20	11:11	Indicated alarm after production of a lot of flakes; after number of non-toolmaking solutions
P14	24/8/20	10:52	"What can I do with the [granite block]?"; after successful flaking via STMU
		11:00	"I could keep cutting with glass but that can't be the meaning of the tasks"
		11:33	"How fast do people realize it's made of glass?"
		11:38	"The big one I cannot move so only use as tool or to make tools." "Stone does not cut so well, so glass is better." "This is the most efficient method, using [glass core] is not effective because it will be marked."
		11:42	Mentioned "comparative study with apes"
P19	8/9/20	10:40	Made a flake and said "Entschuldigung" (EN. "Excuse me"), did not use flake, asked "Can I break this [glass blank]?"
P26	17/9/20	10:24	First time making flakes "Entschuldigung." When leaving after session "I didn't know the entire time it had to do with stone tools."
P28	2/10/20	10:34	"This safety gear is super good, the Neandertals would not have had anything like that"; after 4 toolmaking bouts.
		12:00	Participant mentioned being compared to monkeys.
P30	8/10/20	10:30	"Am I allowed to break the objects?"

Table 8 Breakdown of participants by naivety level and behavioral outcomes. Reported values are the number of participants that fulfill both conditions (specified naivety level and specified behavioral outcome). The values reported in parentheses are those participants that also fulfill the criteria of the cell, but only did so after they had received the Impossible Flake (supplementary text). P6, P21, and P23 all performed knapping techniques in potential toolmaking events before receiving the Impossible Flake, but only P6 made and used cutting tools afterwards. P21 produced potential cutting tools without using them after receiving the Impossible Flake. P23 did so as well, but this was identifiable only from the artefactual outcomes and not via the coding of the video-recorded behavioral data.

Behavioral Outcomes	Naivety levels: Technique-naïve, *n* = 25			Non-technique-naïve, *n* = 3		
	0	*1*	*2*	3	4	
	No knowledge of stone tools or knapping	Heard of stone tools	Seen stone tools	Seen knapping	Hands-on knapping experience	Total
Performed at least one knapping technique (potential toolmaking)	*n* = 2	*n* = 1	*n* = 22	*n* = 1	*n* = 2	*n* = 28
Produced at least one potential cutting tool (confirmed toolmaking)	*n* = 2	*n* = 1	*n* = 19 (*n* = 2)	*n* = 1	*n* = 2	*n* = 25 (*n* = 2)
Produced and used a cutting tool	*n* = 2	*n* = 1	*n* = 19 (*n* = 1)	*n* = 1	*n* = 2	*n* = 25 (*n* = 1)
					Total	*N* = 28

events, 1095 (69.3%) resulted in the fracture of an object and subsequent creation of potentially usable cutting edge (confirmed toolmaking events; Figures 22 and 24). Note that, as what typically occurs during knapping, a single bout sometimes produced more than one detached piece. The combined artifact assemblage at the end of this study consisted of 1599 objects, with flakes accounting for 73.3% of this total (1172 flakes produced from 33 glass cores; Figure 25).

All confirmed toolmaking events for the first totally naive participant P11 (*n* = 36, 100.0%) involved the application of bipolar technique. The preferred approach (used in more than half

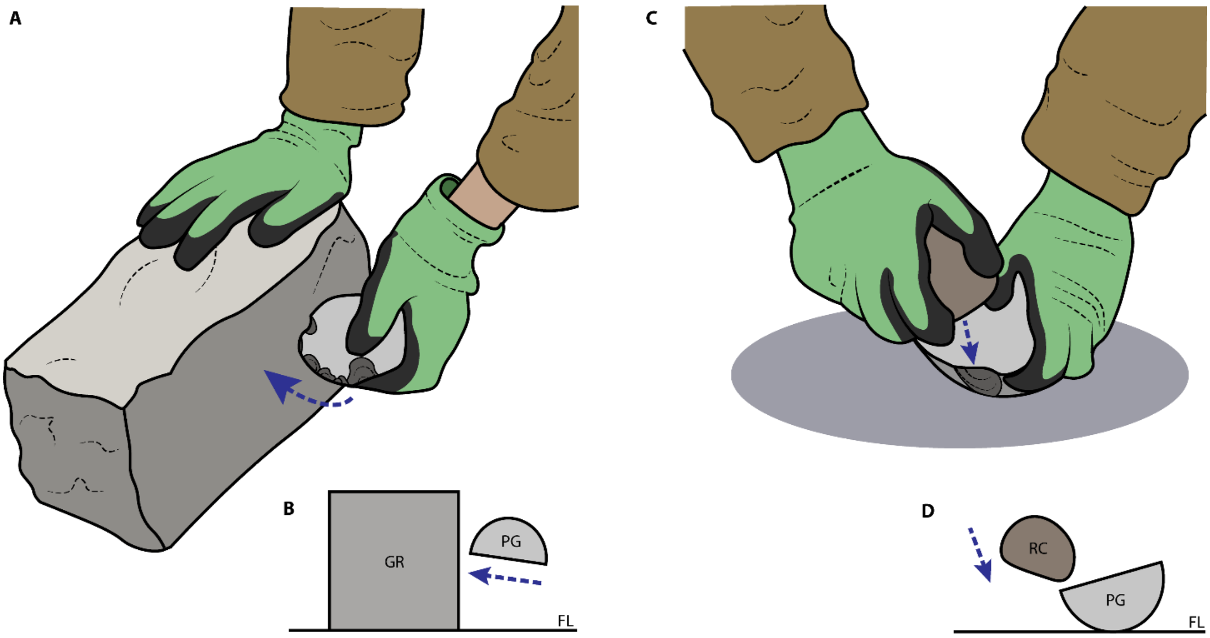


Figure 22 The preferred knapping techniques of the two totally naïve participants (Level 0). (**A** to **B**) Passive hammer technique used by P14 (Trial 45) shown from the perspective of Camera 2 (**A**) and in profile view (**B**). (**C** to **D**) Bipolar technique used by P11 (Trial 23) shown from the perspective of Camera 2 (**C**) and in profile view (**D**). Blue, dashed arrows represent the directionality of force of the active element. The abbreviations in the profile views (**B** and **D**) stand for the following: GR is the granite block, PG is the painted glass hemisphere, RC is the river cobble, and FL is the concrete floor.

of the toolmaking bouts; Figure 22C to D, Figure 24) for P11 was to place the core (glass hemisphere) directly on the concrete floor, stabilize the core with one hand, and then strike it from above with the provided river cobble (as a hammer) using the other hand. When considering potential toolmaking events more broadly, P11 also showed evidence for both passive hammer (n = 1) and freehand (n = 1) techniques. As for products, P11's toolmaking primarily resulted in the production of flakes (n = 35 of N = 41 artifacts in P11's assemblage; Figures 26 and 27).

The first technique used by the other totally naive participant P14 was freehand (n = 1, 1.9%), although nearly all confirmed toolmaking events for P14 (n = 53, 98.1%) involved a variant of passive hammer technique (Figure 22A to B, Figure 24), in which the outer edge of the core (glass hemisphere) was struck against the side of the granite block. Regarding potential toolmaking events, P14 additionally engaged in bipolar technique (n = 1). Just as with P11, P14's toolmaking actions produced mostly flakes (n = 61 of N = 77 artifacts; Figures 26 and 27).

Among all 28 participants, 22 used more than one technique to successfully create sharp edges/cutting tools. All but one participant demonstrated a preference for one technique over

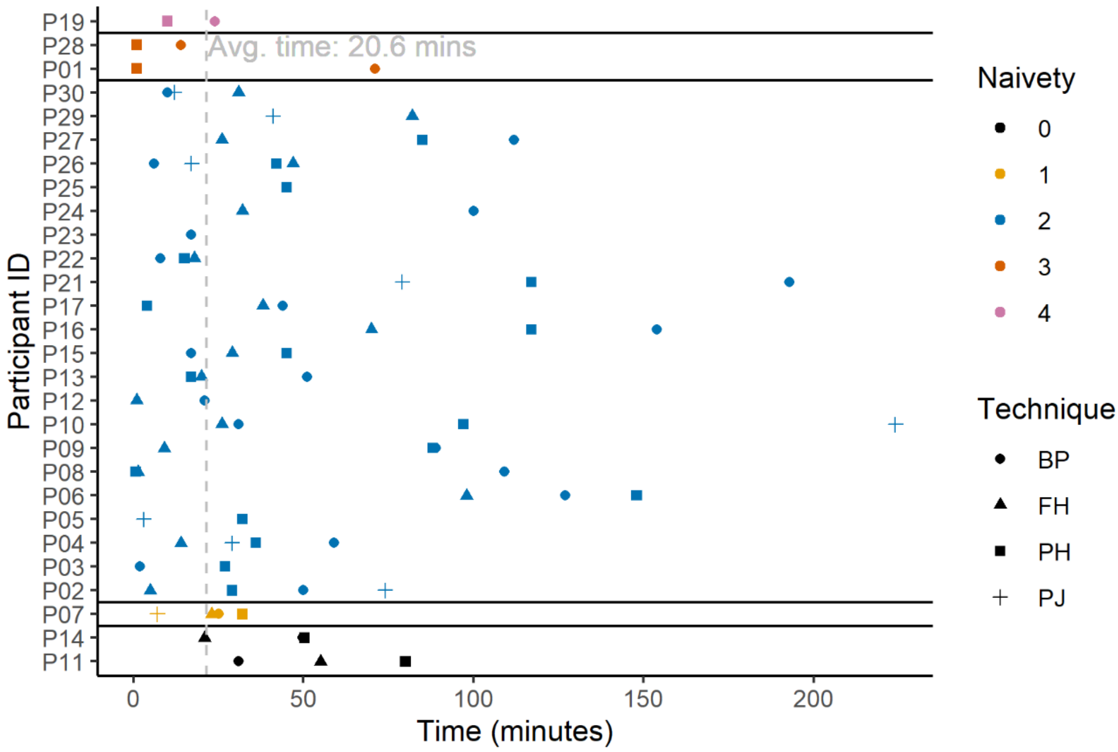


Figure 23 Timeline of first toolmaking innovations by technique. The timing of each innovation of a potential toolmaking technique for each participant (testing time elapsed at the start of the bout, in minutes). Shape of data points corresponds to the technique used in the behavioral bout and color of the data points corresponds to the naivety of the participant based on the post-test questionnaire (see legend). Grey, dashed, vertical line represents the average timing of the first innovation of any technique by each individual.

other techniques. The total number of techniques expressed by individuals did not differ across naivety levels ($X^2 = 5.7528$, df = 4, $p = 0.2184$, Kruskal-Wallis test).

The three early knapping techniques (passive hammer, bipolar, and freehand) reinnovated by the two totally naive participants were also often used by other participants. The fourth coded category, projectile technique, although not shown by either totally naive participant, was developed by several other participants, all of whom were knapping technique naive ($n = 1$ for level 1 and $n = 9$ for level 2; Figure 22).

Across all attempts and across all participants, bipolar technique was the most frequent category of potential toolmaking (574 events, 36.3%), followed by passive hammer (559 events, 35.4%), freehand (410 events, 25.9%), and projectile (37, 2.3%). In terms of individual preference for technique, the preferred technique in potential toolmaking events was most often bipolar technique ($n = 10$, 35.7%), followed by passive hammer ($n = 7$, 25.0%), anvil-oriented (preference for subsequent or alternating use of passive hammer and

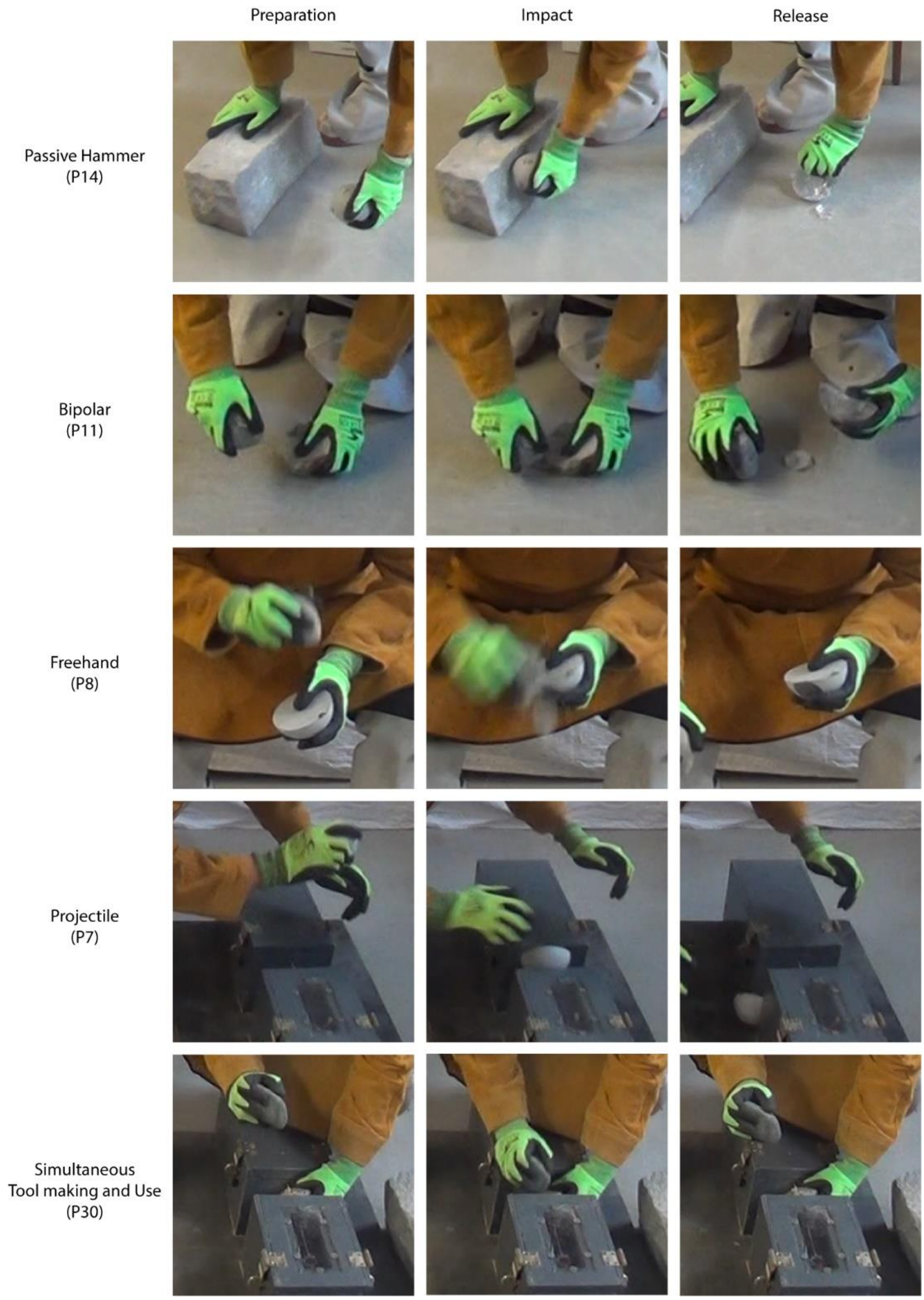


Figure 24 Still frames of tool making technique action sequences. Top row, passive hammer technique (performed by P14, Trial 33). Second row, bipolar technique (performed by P11, Trial 41). Third row, freehand technique (performed by P8, Trial 2). Fourth row, projectile technique (performed by P7, Trial 2; no fracture occurred). Bottom row, simultaneous toolmaking and use, freehand technique (performed by P30, Trial 68). Video Still Credits: William D. Snyder, University of Tübingen.

bipolar technique, *n* = 5, 17.9%) freehand (*n* = 4, 14.3%), and opportunistic (no technique preference, *n* = 2, 7.1%. As far as confirmed toolmaking events, bipolar technique was the most frequent (466 events, 42.6%), followed by passive hammer (398, 36.3%), freehand (226, 20.6%), and projectile (5, 0.5%). Individualized preferences for techniques were distributed between bipolar (*n* = 13, 48.1%), passive hammer (*n* = 10, 37.0%), freehand (*n* = 3, 11.1%), and opportunistic (no technique preference, *n* = 1, 3.7%).

The first instance of potential toolmaking generally occurred within the first hour of testing ($\bar{t}$ = 20 min and 38 s, SD = 25 min and 13 s; Figure 23), with the confirmed first toolmaking also being relatively expedient ($\bar{t}$ = 35 min and 58 s, SD = 56 min and 20 s). Participants typically engaged in numerous noncutting solution attempts (e.g., simple percussion with the cobble against the rope, rubbing a blunt object against the rope, and untwining the fibers of the rope with their fingers) before they showed any potential toolmaking behaviors (mean = 11.7 attempts, SD = 16.7 attempts). This count of attempted noncutting solutions to the puzzle box increases when only considering the first confirmed toolmaking event (mean = 17.9 attempts, SD = 26.1 attempts). Only three individuals (P1, P3, and P28) produced and used a cutting tool as their very first solution to the puzzle box during the test. Some participants attempted noncutting solutions even after at least one initial success in making and using a cutting tool (most notably, P2). One of the two totally naive participants, P11, produced their first cutting tool after 20 min and 52 s of testing time. The second of the two totally naive participants, P14, produced their first cutting tool after 23 min and 23 s of testing time. There was no significant difference between naivety categories in terms of the timing of the first innovated potential toolmaking (X^2 = 5.8528, df = 4, *p* = 0.2104, Kruskal-Wallis test) or of the first confirmed toolmaking (X^2 = 5.3537, df = 4, *p* = 0.2529, Kruskal-Wallis test).

There were 260 behavioral bouts of toolmaking that also *simultaneously* involved actions related to opening the baited puzzle box (simultaneous toolmaking and use), and in 87 such cases (33.5 %), the coder could visibly identify the creation of a cutting tool at the same time that an attempt was made to open the puzzle box. Fourteen participants (P7, P9, P10, P11, P12, P14, P16, P21, P24, P26, P27, P29, and P30; 50.0 %) produced their first cutting tools in such circumstances, after which most – with the exception of one individual using simultaneous methods throughout their session – would eventually prioritize non-

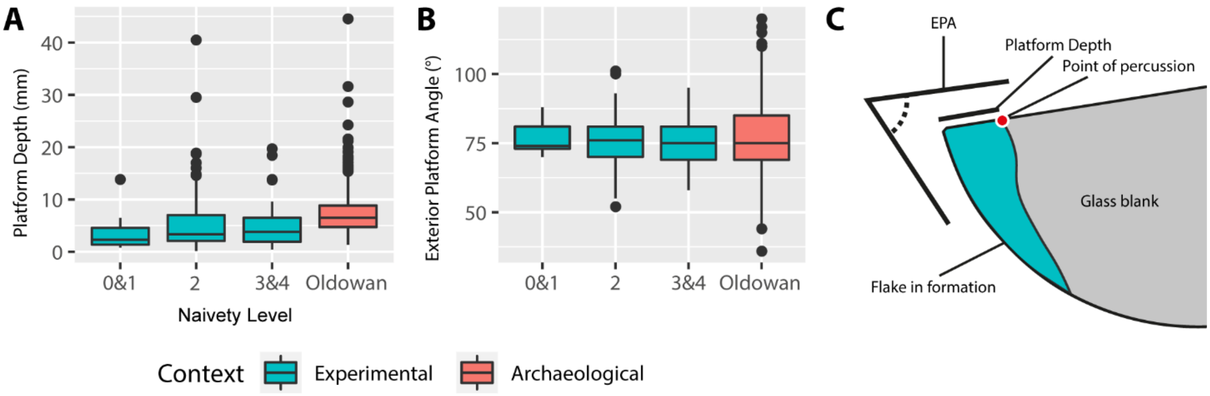


Figure 25 Metric comparison between experimental (*n* = 266) and archaeological flakes (n = 620). (**A**) Depth of platforms on experimental (left) and archaeological (Oldowan; right) flakes. (**B**) Exterior platform angle on experimental (left) and archaeological (Oldowan; left) flakes. (**C**) Schematic showing the formation of a flake by percussive knapping and the relationship to the aforementioned metrics.

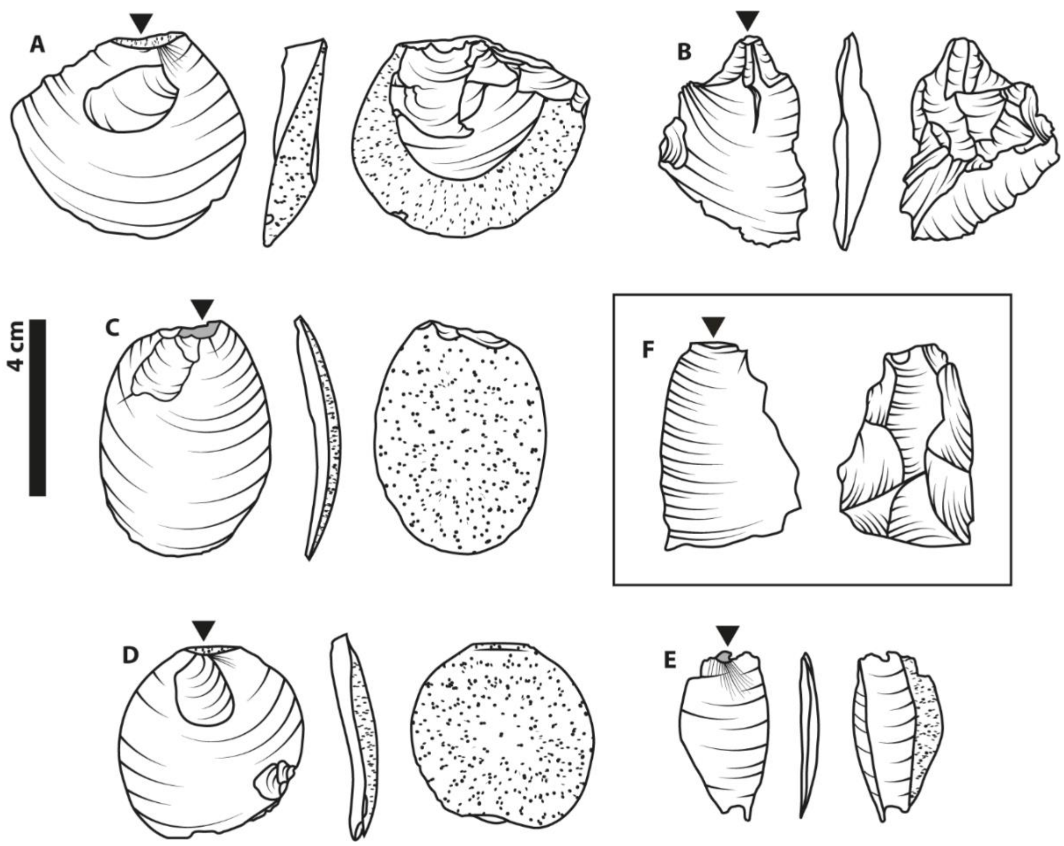


Figure 26 Technological illustrations of flakes. All experimental flakes depicted here were made by individuals who had never before seen the process of stone toolmaking and were used successfully as cutting tools. Experimental artefacts are shown here in ventral, lateral, and dorsal perspective. Flakes are shown in platform-up orientation, with black triangles pointed at the approximate location of the point of percussion. Light grey areas indicate crushing events while dotted areas indicate cortex (the original painted surface of the glass hemispheres). Participant ID (e.g., P11), participant naivety level (e.g., L0 for totally naive), and the knapping technique that was used to produce the flake are as follows: P11, L0, bipolar (**A**); P14, L0, passive hammer (**B**); P7, L1, passive hammer (**C**); P8, L2, freehand (**D**); and P15, L2, projectile (**E**). A pre-modern hominin-produced flake (**F**) from the site of Gona, Ethiopia (2.6 - 2.5 Ma; Semaw, 2006) serves as just one example of a vast diversity of flakes known from the Oldowan and is shown in the box, center-right.

Table 9 Unsuccessful and excluded participants. *QUP* is the Questionnaire for Unsuccessful Participants, which was given in two parts to those participants that failed to make and use cutting tools. *PSQ* is the questionnaire titled "Additional questions in case of a partial solution", which was given to participants that created cutting tools but failed to use them to open the puzzle box. Here, we report the full answers, translated from German to English (see Appendix III). Participants highlighted in orange were those whose data were excluded from the analyzed datasets.

Individual	Performance	Result	QUP Part 1	QUP Part 2	PSQ
P2	Produced an angular fragment from granite and used it for cutting the rope, but subsequently resorted to other apparatus solution techniques	Failure to reproduce the initial toolmaking success, participant given QUP	1: "If humans in our century are capable of using stone tools, or rather can come to the same approaches as in the time that these objects were the only tools" 2: "Too many own thoughts, if a new idea is again rubbing (or instead maybe cutting). Not attempted, if I myself thought that it is really nothing new"	1: "Yes" 2: "Yes, the thought that I should find a new approach. Rather thought about physics than on the use of objects. The thought that everything I used would be taken away, I thought ultimately that I must also come to a solution without the objects"	n.a.
P6	Provided impossible flake, subsequently produced and used glass cutting tools	Treated as success, so not given QUP or PSQ	n.a.	n.a.	n.a.
P21	Provided impossible flake, produced viable cutting tools but did not use them to open tendon box	Failure to use participant-created cutting tools, participant given QUP and PSQ	1: "The behavior of humans by the solving of difficult tasks." 2: "Creativity, i.e. to get the door open with many different options. Possibly, that there are still other solutions, e.g. to not sever the rope"	1: "Yes" 2: "Infrequent use of stones, thus it was difficult to spontaneously make a "tool""	Realized that sharp-edged objects (i.e. glass flakes) had been produced but did not used them because "[they] already had used many options to cut apart the rope"
P23	Provided impossible flake, produced viable cutting tools but did not use them to open tendon box	Failure to use participant-created cutting tools, participant given QUP and PSQ	1: "How people deal with unfamiliar situations, for which they may have a hard time finding solutions" 2: "Inexperience with practical work and insecurity due to perplexity"	1: "Yes" 2: "In the course of schooling learned things were often repeated or rather applied, yet the problem solving competence was more rarely expected"	Did not realize that sharp-edged objects (i.e. glass flakes) had been produced and felt as if "[their] ideas were exhausted and also through further reflection, no new ideas arose"
P18	Safety concern	Session ended, data excluded	n.a.	n.a	n.a
P20	Experimenter error	Data excluded	n.a.	n.a	n.a

simultaneous methods (i.e., toolmaking followed by use in separate, successive behavioral bouts).

Three out of the 28 participants did not produce and/or use cutting tools within the first 2 hours of the testing session (Table 9). Therefore, according to the predesigned protocol, at this point these participants received the "Impossible Flake". All three participants used the impossible flake as a tool to then access the reward. Afterwards, one of these three participants (P6) produced and used glass cutting edge, while the two remaining participants of these three (P21 and P23) produced glass cutting edge but then failed to use it as a tool to open the tendon box. Of the latter two, P21 reported in the PSQ that he recognized that a cutting edge was made but felt there were other solutions worth attempting, while P23 reported in the PSQ that she failed to recognize that she had created any useable glass cutting edge.

Another participant (P2) produced an angular fragment from the granite brick and used it to sever the rope but was not able to repeat this or other techniques and instead opted to pursue other means of opening the tendon box.

Participants across naivety categories did not differ in terms of the number of flakes produced ($X^2 = 3.0340$, df = 4, $p = 0.5522$, Kruskal-Wallis test). Experimental flakes had statistically shallower platforms [platform depth (PD)] than Oldowan flakes ($X^2 = 111.96$, df = 3, $p = 2.2 \times 10^{-16}$, Kruskal-Wallis test; Figure 25A). All exterior platform angles (EPAs) measured on experimental flakes (range of 52° to 101°) fell within the expected range for Oldowan flakes (Figure 25B) (Režek et al., 2018). EPA was not significantly different between naivety groups and between naivety groups and Oldowan assemblages ($X^2 = 0.53688$, df = 3, $p = 0.9107$, Kruskal-Wallis test).

Patterns of removals (Figure 25C) on the cores themselves varied across individuals (Figure 27). The most frequently appearing morphology were multifacial cores (*n* = 9; Figure 27C), which were heavily reduced without clear evidence of systematic exploitation of specific knapping platforms. After this, the core type with the most frequent and identifiable pattern of removals were radial cores (*n* = 8; Figure 27A). These cores involved the exploitation of the rounded surface of the hemisphere as a striking platform, causing mostly removals from the flat face that are directed at the midpoint of the circle. Knapping of a single exploitation surface or on two adjacent exploitation surfaces on one edge of the blank produced cores similar to unifacial (*n* = 5; Figure 27B) and bifacial (*n* = 2; Figure 27D) choppers, as well as heavy-duty (n = 2) scrapers. One Karari-style scraper (n = 1) (Harris & Isaac, 1976) was also produced by the complete exploitation and removal of the rounded surface of the blank.

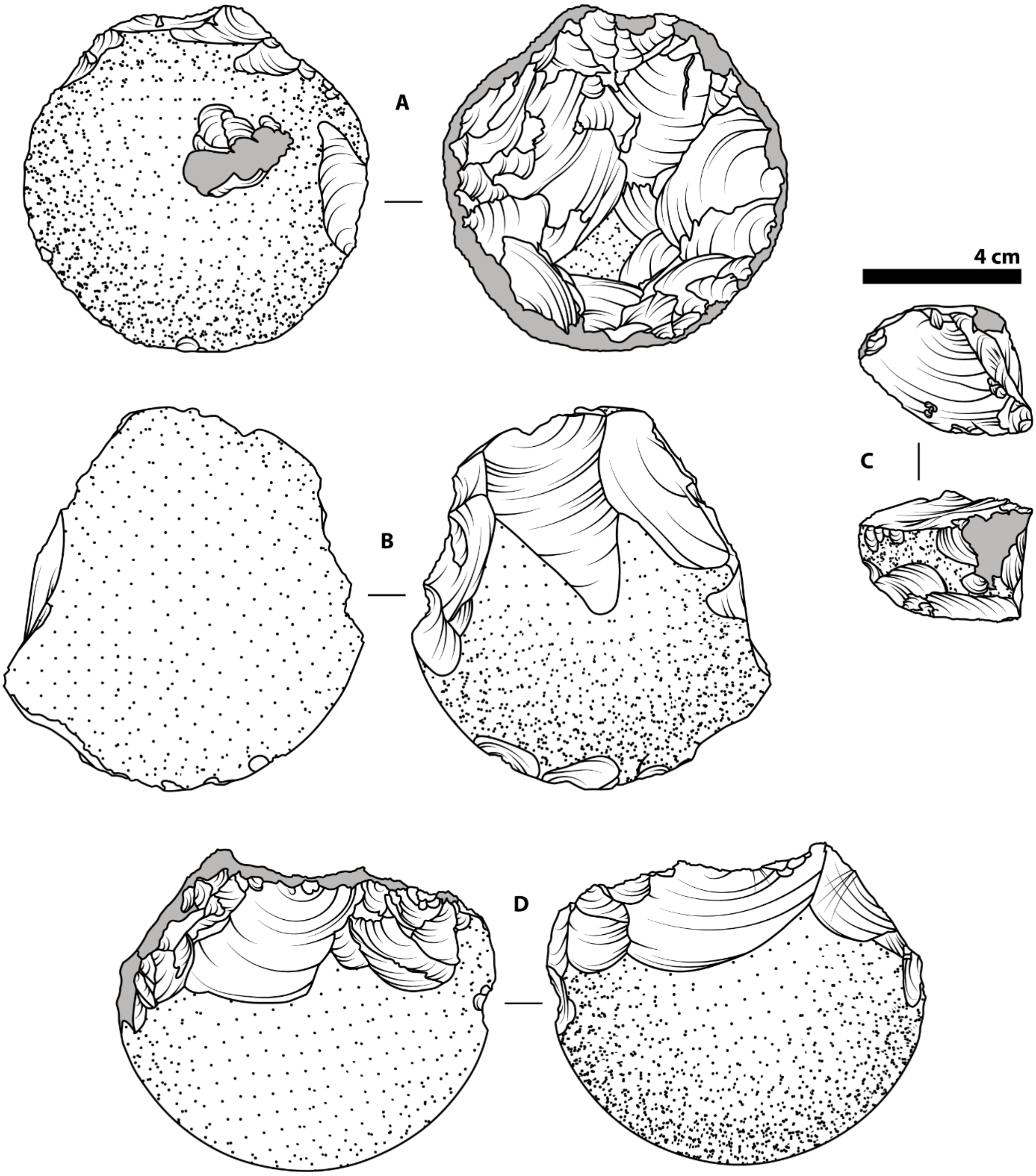


Figure 27 Representative cores from this study. Illustrations based on a selection of cores produced by the participants in this study. As with the Oldowan artefacts (Fig. 1), represented here are a variety of core types including: a radial core produced by P14 (**A**), a unifacial core or "chopper" produced by P7 (**B**), a multifacial core produced by P17 (**C**), and a bifacial core produced by P19 (**D**).

Expanded results

Analyses of the behavioral and artifactual outcomes for this study are ongoing. The following are preliminary results from said ongoing analyses.

Detached pieces

We note that many participants (*n* = 13) also made sharp stone tools from non-glass materials, i.e., from the river cobbles and the granite block. Some continued to attempt removals from the river cobbles even after they had already successfully produced glass cutting tools.

Debitage pieces of several participants (*n* = 12) showed signs of edge modification. On average, 20.66 % (SD = 15.45) of the edge of the flake or flake fragment indicated modification by small removals. Most of this evidence of flake edge modification could be corroborated by the behavior coding data (*n* = 5 individuals overlapped, including the most frequent agent of flake modification, P22).

Comparisons between experimental and archaeological flakes (Table 10 and Figure 28) revealed statistically significant metrical differences between the two samples in nearly all measures analyzed (including mass, maximum dimension, width, thickness, and platform depth, but not EPA). Experimental flakes – i.e., those created by the novices from our study –

Table 10 Basic metric attributes of flakes from Oldowan sites (Režek et al., 2018) and from the experiment (Snyder et al., 2022).

Metric attribute		Archaeological	Experimental
Mass (g)	Mean	31.34	5.30
	Std. dev.	39.00	11.22
Maximum dimension (mm)	Mean	50.95	30.83
	Std. dev.	42.40	13.04
Width (mm)	Mean	40.87	18.21
	Std. dev.	31.37	9.87
Thickness (mm)	Mean	13.15	4.86
	Std. dev.	8.76	4.17
Platform depth (mm)	Mean	7.31	4.99
	Std. dev.	4.01	4.73
EPA (°)	Mean	76.55	13.05
	Std. dev.	75.56	8.60

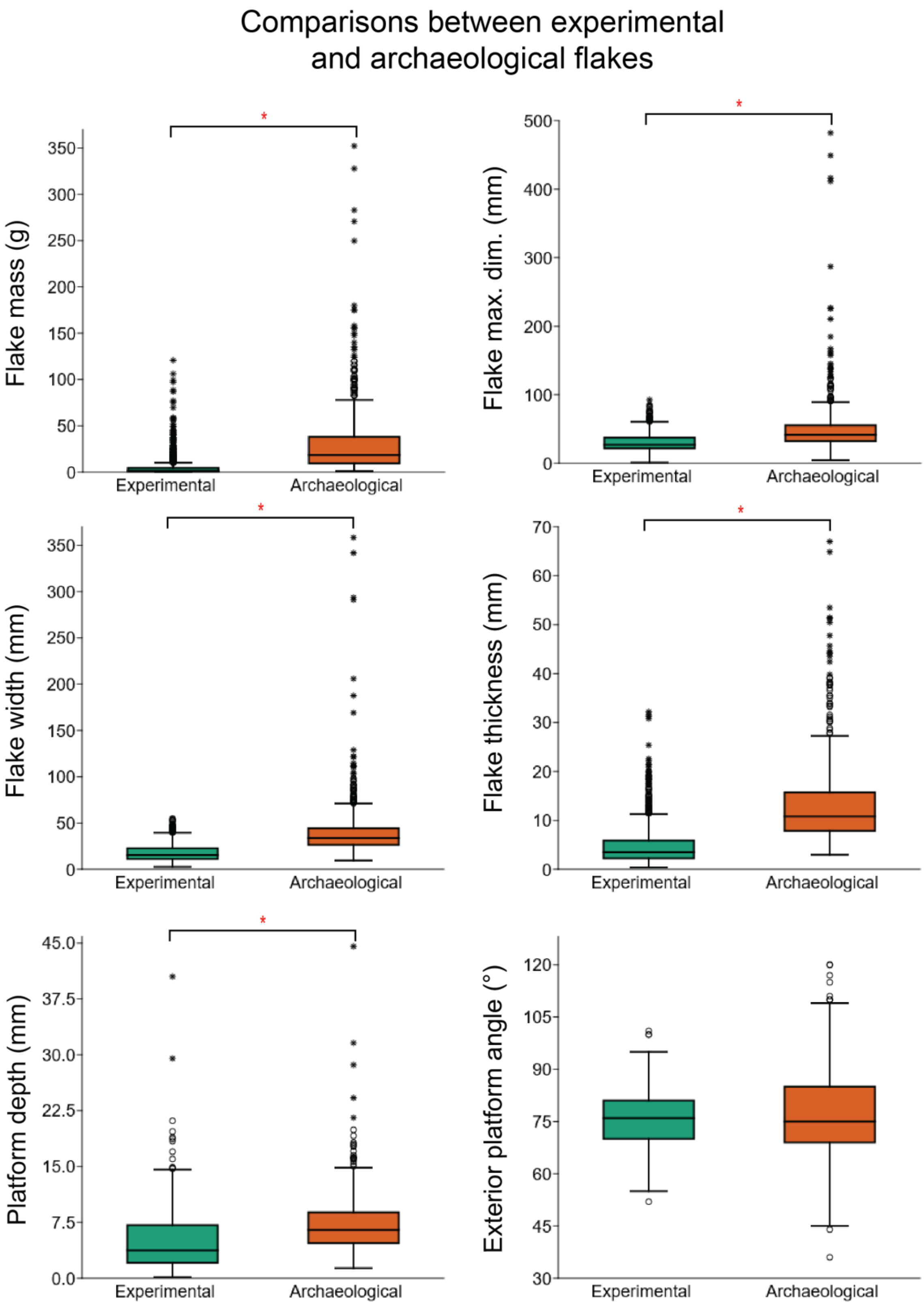


Figure 28 Box plot visualizations of the metric comparisons between flakes from experimental and archaeological contexts. Red asterisks mark differences that are statistically significant.

were, on average, lighter (Z = 27.79, p < 0.0001, Wilcoxon rank-sum test), smaller (t = -11.53, df = 681.6, p < 0.0001, Student's t-test), narrower (t = -17.53, df = 684.5, p < 0.0001, Student's t-test), and thinner (t = -22.26, df = 770.50, p < 0.0001, Student's t-test) than the flakes from actual Oldowan assemblages. Experimental flakes also possessed shallower platforms (t = -6.97, df = 433.81, p < 0.0001, Student's t-test). To the contrary, there was no statistically significant difference between the EPAs of the two sets of flakes (t = -1.31, df = 738.30, p = 0.1897, Student's t-test), with the EPAs of the experimental flakes falling fully within the range shown by flakes from Oldowan assemblages.

Glass cores

On glass cores, there is a lack of truly invasive flaking (e.g., for centripetal-type cores, there is always cortex remaining at the center of the horizontal plane) and effective volume management. Many cores show indication of battering, a consequence of their inability to effectively exploit suitable knapping surfaces and persistence in attempting fracture unsuitable surfaces. This is further supported by the frequency of step and hinge fractures on the surface of cores (n = 30 cores). All cores with >5 removals show indications of step or hinge terminations on their surfaces, though the ratio of step/hinge terminations to overall scar count varies from core to core (Figure 29). Overall, just under half of all flake scars (48

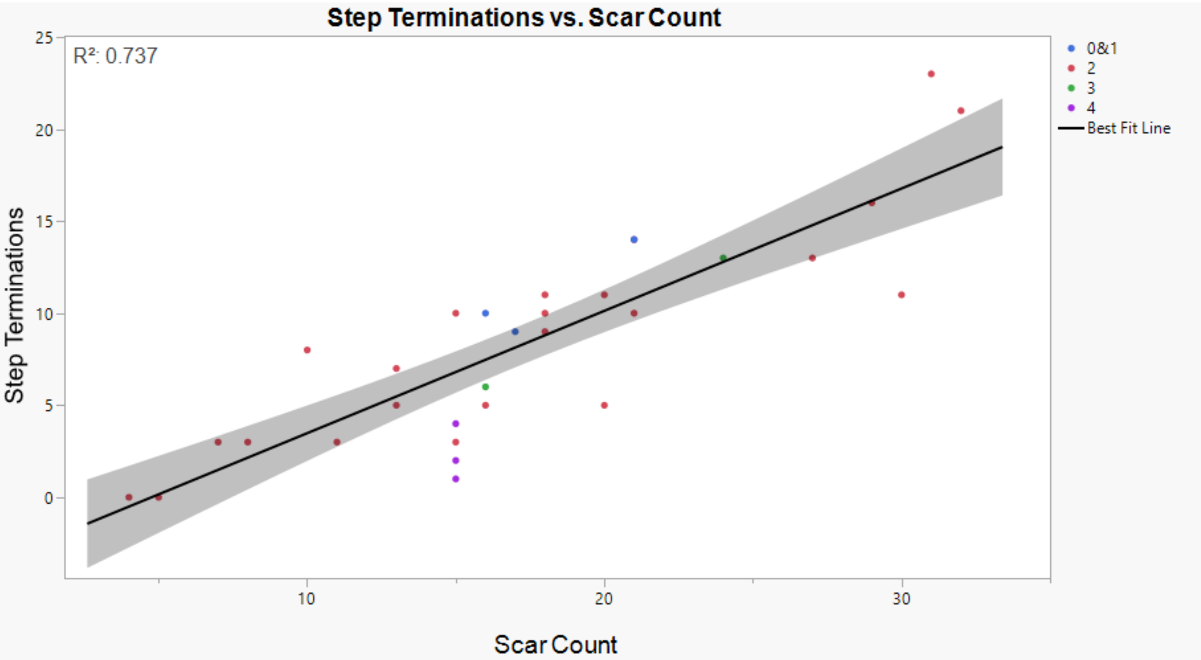


Figure 29 Step terminations plotted against scar count, showing the linear relationship between the two variables.

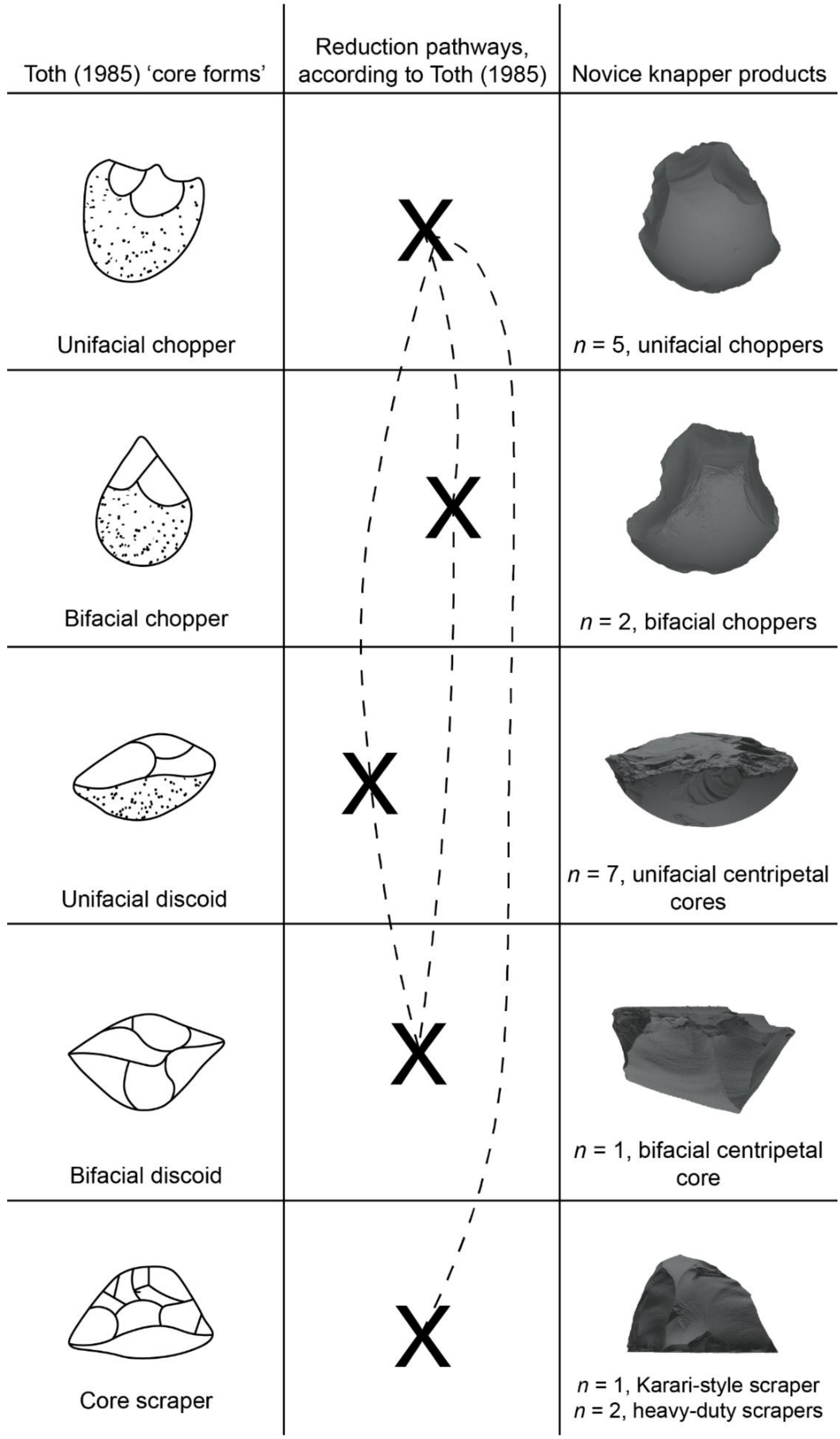


Figure 30 Predicted 'core forms' (left) and reduction pathways from hemispherical blanks according to Toth (1985, fig. 5). Representative cores from the naïve knapper experiments (right) are given for each of the 'core forms'.

Table 11 Attributes of individual blanks from the experiment (Snyder et al., 2022), in order of the participant that produced them. Metric attributes include scar counts, exploitation surface counts, interaction surface counts, and the percentage of blank left unexploited. Technological assessment of the cores was both typological (using classical types; Leakey, 1971) and reduction sequence-based (according to methodology of de la Torre, 2011).

Part.	Blank	Scar Count	Exploit. surfaces	Interact. surfaces	% of blank unexploited	Type (Leakey, 1971)	Reduction (de la Torre, 2011)
Z1	C6	24	4	3	78.35	Discoid	UC
Z3	C1	15	2	0	77.5	Unifacial chopper	USP
Z4	C9	21	3	1	79.9	Heavy-duty scraper	BAP
Z5	C10	20	3	1	65.3	Bifacial chopper	BSP
Z7	C12	16	3	3	73.7	Bifacial chopper	BSP
Z8	C13	13	1	0	22.9	Core scraper	UAUT
Z9	C14	18	5	3	21.5	Multifacial	Multifacial
Z10	C15	16	5	3	42.3	Unifacial chopper	USP
Z11	C16	17	3	1	74.9	Unifacial chopper	USP
Z12	C17	10	>3	1	8.3	Multifacial	Multifacial
Z12	C4	7	3	1	50.3	Scraper	UAU1
Z13	C18	21	5	3	17.2	Multifacial	Multifacial
Z14	C19	21	4	2	58.3	Discoid	UC or UP
Z15	C20	7	4	0	14.0	Multifacial	Multifacial
Z16	C21	13	7	2	21.4	Multifacial	Multifacial
Z17	C22	15	7	1	8.6	Multifacial	Multifacial
Z17	C5	11	9	2	10.0	Multifacial	Multifacial
Z17	C23	18	4	3	78.4	Discoid	UABI
Z19	C25	15	9	6	4.7	Multifacial	Multifacial
Z19	C25	15	15	2	11.4	Multifacial	Multifacial
Z19	C26	15	1	0	80.0	Unifacial chopper	USP
Z22	C29	20	13	1	19.9	Multifacial	Multifacial
Z24	C31	18	4	2	54.2	Unifacial chopper	USP
Z25	C32	31	3	3	66.3	Discoid	UC or UP
Z26	C33	27	6	5	80.2	Discoid	UC or UP
Z27	C34	30	3	3	84.3	Discoid	UC or UP
Z28	C35	16	3	4	15.1	Discoid	BP
Z29	C36	32	2	2	47.3	Heavy-duty scraper	BAP
Z30	C37	29	2	2	72.8	Discoid	UC or UP

%) end in step or hinge terminations, with an average of 17.1 scars per core and an average of 8.2 step/hinge terminations per core. There is a linear relationship between total scar count and step terminations (across cores from all participants, regardless of naivety groups), suggesting a correlation between the stage or intensity of core reduction and the count of step terminations (R^2 = 0.737).

The total count of scars and the proportion of step terminations to flake scars is comparatively higher than for cores from Oldowan assemblages. For instance, at the early Oldowan site BD1, 31% of flake scars end in step terminations but a majority of cores have only one or two flake scars (Braun et al., 2019). Relative high scar counts could be related to non-invasive flaking (e.g., related to knapper skill and effectiveness) and raw material properties, but also might be related to the stage of reduction (see Braun et al., 2005).

Overall, there is a large amount of variation in terms of both morphological types and reduction strategies (Figure 30 and Table 11). Using Toth's (1985) observations of initial forms and probable end products as a prediction, hemispherical blanks should lead to a particular range of products, including choppers, discoids, and scrapers. All of these were represented in the assemblages produced by the naïve knappers of this study.

Multifacial cores

The most frequently appearing category of cores were irregular *multifacial* cores (n = 9 cores; Z9.C14, Z12.C17, Z13.C18, Z15.C20, Z16.C21, Z17.C22, Z17.C5, Z19.C25, Z22.C29).

Multifacial cores (Figure 31) were the outcomes of intensive reductive of the initial blank, with very little remaining blank mass. There was little evidence of targeted systematic exploitation of specific knapping platforms, and instead, the multifacial cores are a good indicator of opportunistic, least effort flaking of any available platforms as the cores become increasingly reduced.

The latter assertion is supported by differences in reduction intensity (based on the percentage of unexploited blank mass) between cores of the different types. The overall differences were statistically significant ($X^2 = 18.7369$, df = 5, $p = 0.0022$, Kruskal-Wallis test). Pairwise differences revealed no differences between any of the other core types with one another, but all other core types had significantly more unexploited blank mass than the multifacial cores (multifacial versus unifacial choppers: $Z = 3.00062$, $p = 0.0027$, Wilcoxon rank-sum test; multifacial versus core scrapers: $Z = -2.04093$, $p = 0.0413$, Wilcoxon rank-sum test; multifacial versus bifacial choppers: $Z = -2.04093$, $p = 0.0413$, Wilcoxon rank-sum test; multifacial versus heavy-duty scrapers: $Z = -2.04093$, $p = 0.0413$, Wilcoxon rank-sum test; and multifacial versus discoid: $Z = -3.15426$, $p = 0.0016$, Wilcoxon rank-sum test). Multifacial cores were, on average, 13.7 % of the original blank mass, compared to 36.6% for core-scrapers, 63.6 % for heavy-duty scrapers, 65.8 % for unifacial choppers, 66.7 % for discoids, and 69.5 % for bifacial choppers.

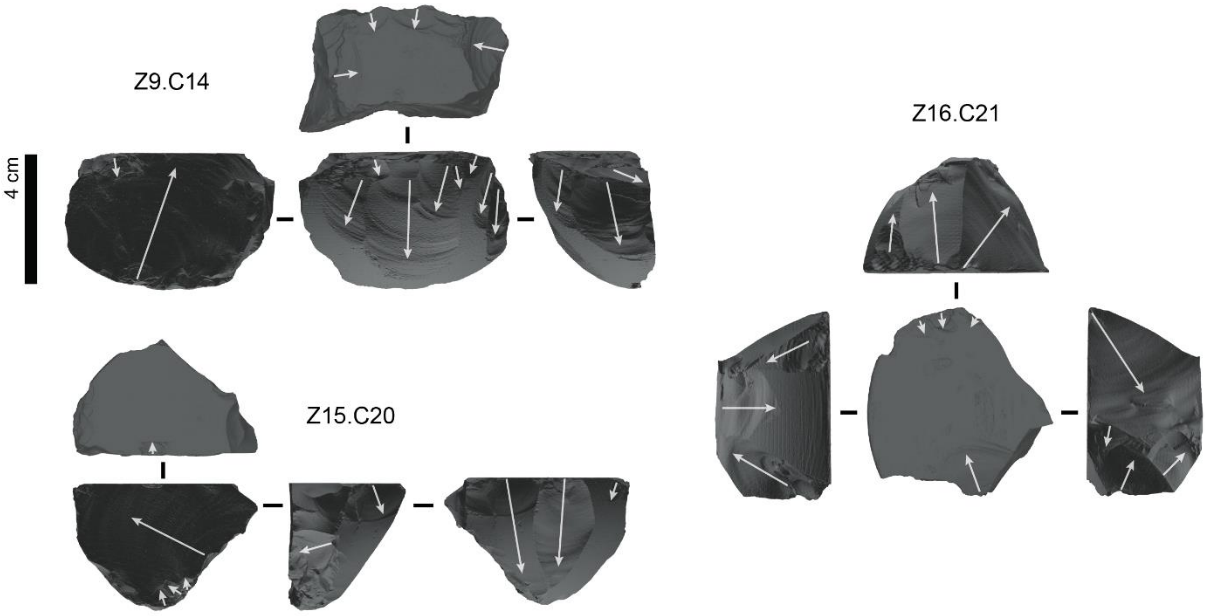


Figure 31 Examples of multifacial cores produced by experiments during the study.

Similar analysis involving total scar counts on cores was not as clear-cut. The overall differences were also statistically significant significant (X^2 = 17.4254, df = 5, *p* = 0.0038, Kruskal-Wallis test), but only two of the pairwise differences were statistically significant (unifacial chopper versus discoid: *Z* = -2.42537, *p* = 0.0153; multifacial versus discoid: *Z* = -2.98581, *p* = 0.00278). Two other pairwise differences were nearly statistically significant (discoid versus core-scraper: *Z* = 1.95837, *p* = 0.0502; multifacial versus heavy-duty scraper: *Z* = -1.95064, *p* = 0.0511). Multifacial cores had the second lowest mean scar count (mean = 14.5 scars per core), after core-scrapers (mean = 10 scars per core).

Centripetal cores

The next best represented category of glass cores were centripetal cores (*n* = 8; Figure 32).

Again, this would likely represent opportunistic utilization of available knappable edges, with the outer edge of the hemisphere serving as an ideal working platform (cf. Delagnes & Roche, 2005). Typically, removals come primarily from the horizontal plane, but often there are some loci where the horizontal surface was used (perhaps inadvertently) as a platform to remove flakes from the rounded plane of the blank. However, these removals are insufficient in consistency to characterize the reduction sequences as bifacial. As such, from these assemblages, most of these cores were exploited using *unifacial centripetal* (*UC*) or *unifacial peripheral* (*UP*) flaking patterns (*n* = 6; Figure 32A). The distinction between the two refers to the directionality of the flake scars, with unifacial peripheral perhaps being a better characterization for cases where the direction of flaking was not completely pointed towards

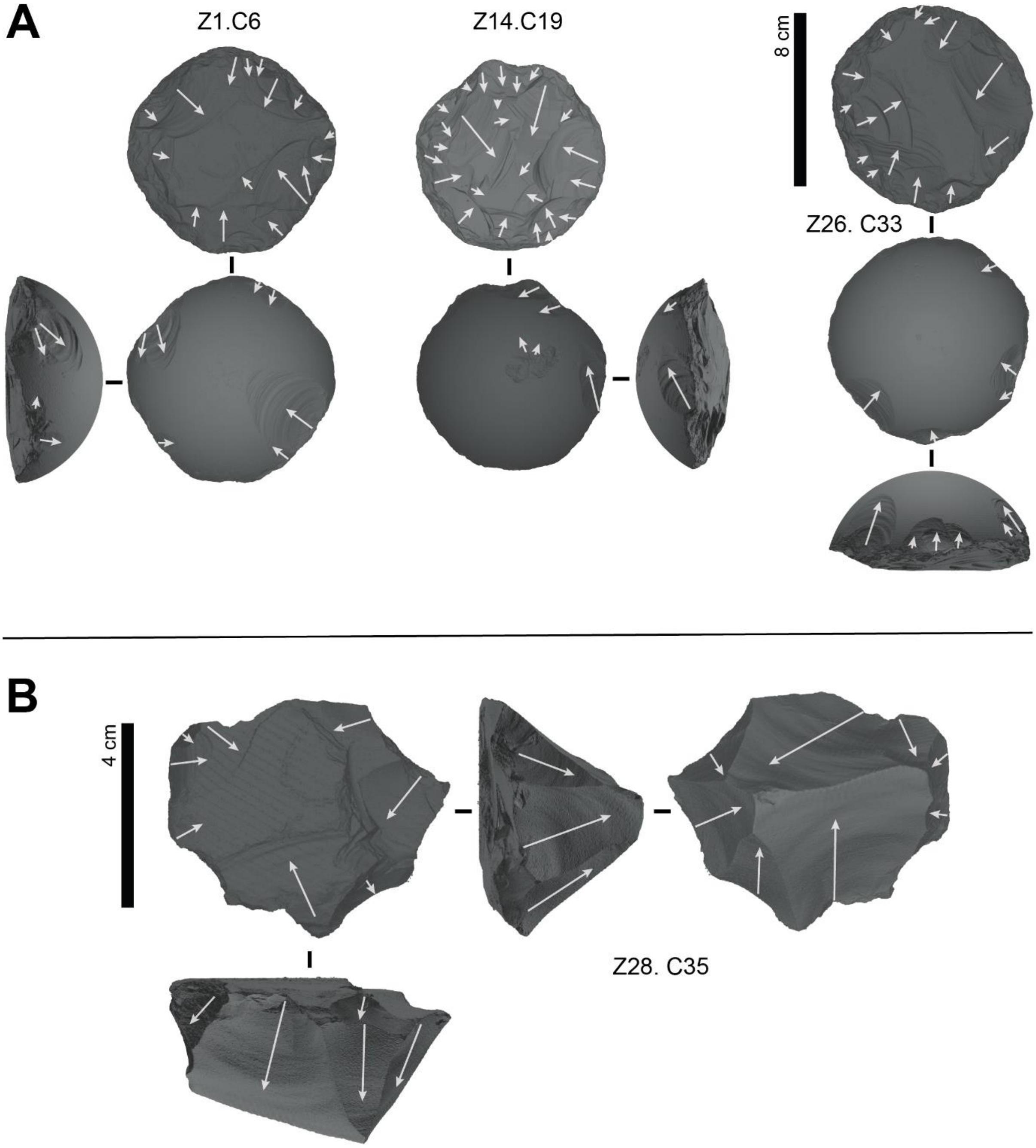


Figure 32 Examples of centripetal cores produced by participants during the study, including unifacial centripetal cores (A) and a bifacial centripetal core (B).

the very center of the flat, horizontal surface of the core. Cores qualifying as having UC or UP reduction patterns include Z1.C6, Z14.C19, Z25.C32, Z26.C33, Z27.C34, and Z30.C37. These cores bear superficial similarities to cores from the earlier Oldowan at Koobi Fora ('unifacial discoid'; Toth, 1985) and Gona ('partial discoid'; Semaw, 2006).

Another core was this group could be characterized as *unifacial abrupt bidirectional exploitation* (*UABI*). This core (Z17.C23) primarily indicates exploitation from the outside edge of the horizontal plane, but with some removals using the horizontal plane as a platform.

Removals from the horizontal plane are at two primary exploitation surfaces that are directionally opposed, hence *bidirectional* exploitation.

There was one case of *bifacial peripheral* knapping (*BP*), Z28.C35 (Figure 32B). The full circumference of core has been exploited on both (i.e., horizontal and rounded) planes. Since removals from the horizontal plane were not consistently directed toward the center of the circle of the horizontal plane, the removals are best considered peripheral in nature. Meanwhile, the full cortex of the rounded plane surface has been removed.

Bifacially-reduced centripetal cores (otherwise, 'bifacial discoids'; e.g., Toth, 1985) are a relatively frequent occurrence in Mode 1 assemblages, especially for the Developed Oldowan and early Acheulean. This includes artifacts from the Oldowan site of Peninj, Tanzania at around 1.6-1.4 million years ago (de la Torre et al., 2003), the Developed Oldowan/early Acheulean site of Gadeb, Ethiopia at around 1.45 million years ago (de la Torre, 2011), the early Acheulean site of Rietputs, South Africa at around 1.3 Million years (Leader et al., 2018), and the *Homo floresiensis* type-site of Liang Bua, Indonesia at around 95 ka (Moore & Brumm, 2009; Moore et al. 2009). Notably, centripetal cores from these later contexts are typically bifacially reduced but also show evidence of hierarchical flaking strategies that are absent from the centripetal cores in this study (*bifacial hierarchical centripetal exploitation* a la de la Torre, 2011). We would not characterize the occasional removals from the rounded surface of the hemisphere on unifacial centripetal cores as being hierarchical but rather infrequent and likely incidental.

Unifacial 'choppers' and 'scrapers'

After the centripetal cores, the next largest contingent were unifacial choppers and scrapers (*n* = 7; Figure 33).

Included in this group were unifacial choppers (*n* =5). In these cases, the horizontal platform was used as a knapping platform, with most removals from the same exploitation surface on the rounded aspect of the hemisphere, thus characterized by *unifacial simple partial exploitation* (*USP*). Cores exhibiting this pattern of removals included Z3.C1, Z10.C15, Z11.C16, Z19.C26, and Z24.C31.

The core Z8.C13 (*n* = 1) had all flakes removed using horizontal plane as a platform, while the original surface of the rounded plane had been totally removed, indicating *unifacial*

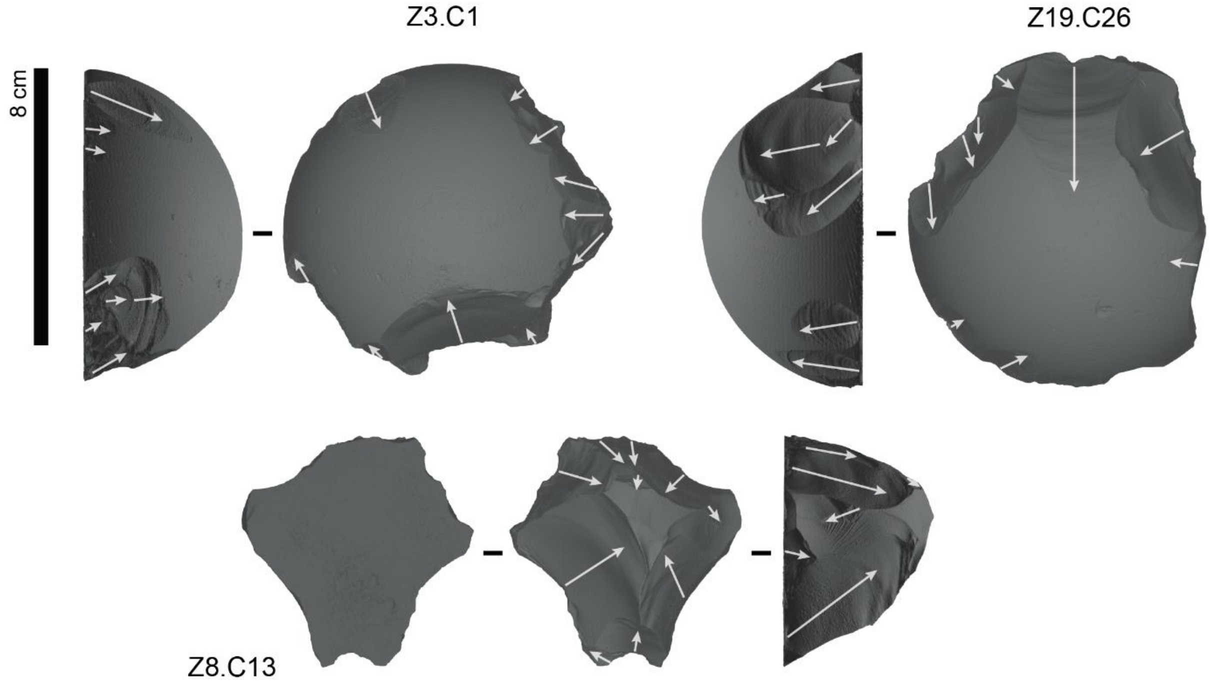


Figure 33 Examples of unifacial choppers and unifacial scraper produced by participants during the study.

abrupt unidirectional total exploitation (*UAUT*). We have typed this core as a 'Karari-style' scraper (Harris and Isaac, 1976).

Another core-scraper (*n* = 1), Z12.C4, is interpreted to have *unidirectional abrupt unifacial* (*UAU1*) exploitation, with most removals coming off of the rounded surface and the horizontal surface being used as a knapping platform. The full circumference of core was not exploited.

Bifacial 'choppers' and 'scrapers'

Bifacial choppers (n = 2) displayed *bifacial simple partial exploitation* (*BSP*), whereby exploitation of the core was almost exclusively from two adjacent striking platform forming a bifacial edge, though there are occasional removals from other exploitation surfaces. Bifacial choppers (Figure 34) include cores Z5.C10 and Z7.C12.

Further bifacially worked cores were heavy-duty scrapers (*n* = 2; Z4.C9 and Z29.C36). These were produced by *bifacial abrupt partial exploitation* (*BAP*). As with BSP, BAP exploitation involved the flaking of the core at two adjacent exploitation surfaces (i.e., at a single interaction surface), meeting each other abruptly, i.e., at an oblique angle (or rather, an angle that is noticeably more oblique that would characterize BSP).

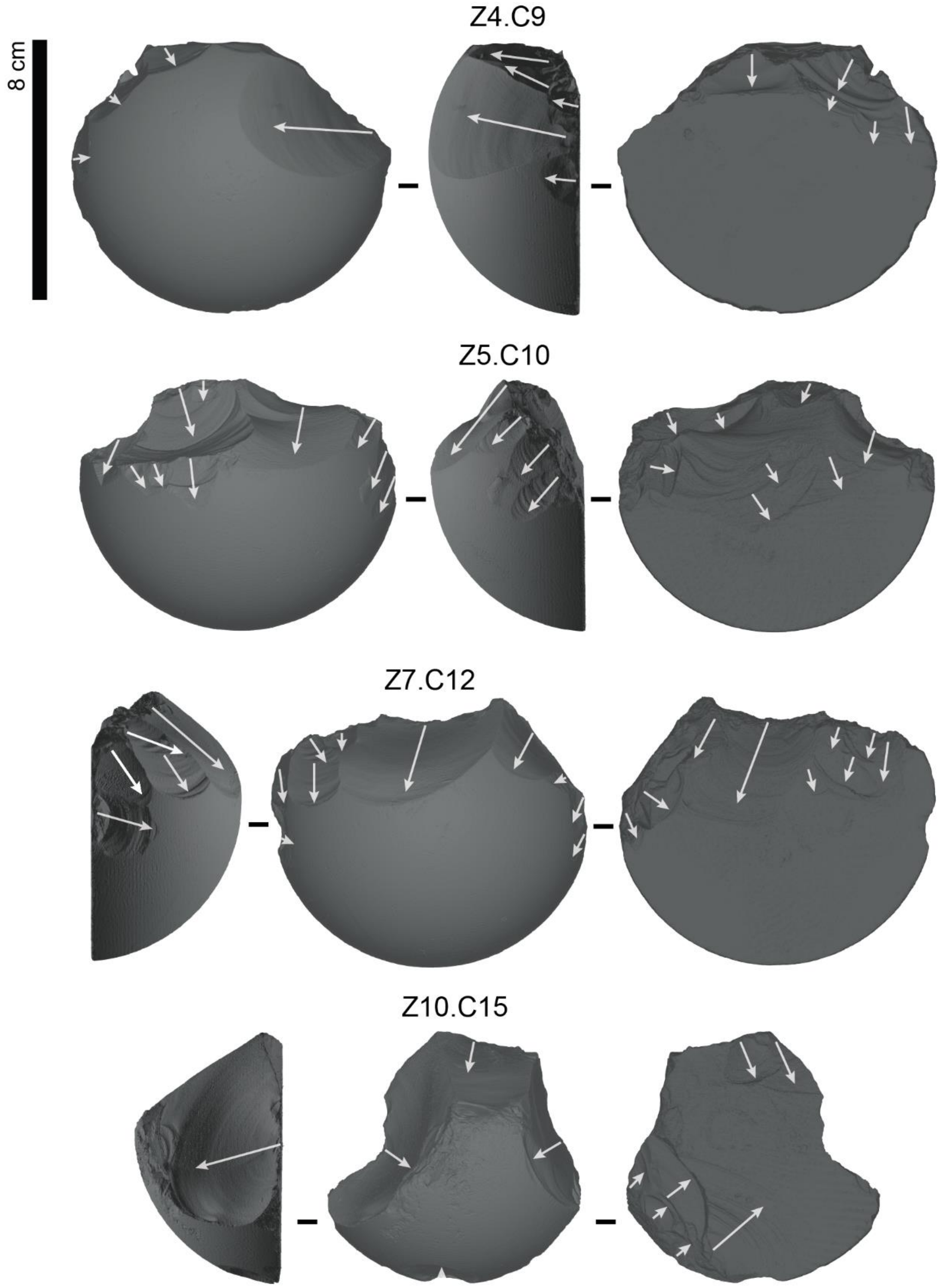


Figure 34 Examples of bifacial choppers and scrapers produced by participants during the study.

Additional chopper and scraper cores might have been similarly classified as products of BSP or BAP exploitation, but removals from a second exploitation surface on the opposite plane were too infrequent and also nonadjacent (therefore, not composing an identifiable exploitation surface). It is unclear if further knapping by the participants would have led to the conversion of unifacial choppers and scrapers into bifacial cores (or into some other category; see Toth, 1985; Moore & Perston, 2016; see also Figure 30).

Protobiface / handaxe

In the course of reducing the glass blank C22 (via a variant of passive hammer technique), participant P17 produced a form with bifacially-worked edges and bilateral symmetry (Figure 35). On this basis, it was initially judged by the experimenter (W.D.S.) as a noteworthy enough product to be photographed, given the minimal resemblance to a biface (Acheulean or Mode 2 technology; Ambrose, 2001; Leakey, 1971; Shea, 2010). The core was temporarily taken to a separate room and photographed (with scale). An exception was not made, however, in terms of the regular core exhaustion rule for the study, so the bifacial specimen was returned to the participant, after which further removals were made from its surface, until the remaining part of the core displayed a multifacial reduction pattern. An attempted refit was performed but was unsuccessful. Typing and attribute analysis were conducted on the exclusive basis of the photographs taken during the test.

Maximum dimension of the ‘unintentional’ biface is 6.60 cm (Figure 35A). This is smaller than the sizes (>10 cm in length; Shea, 2010; or >8 cm in length; Beyin et al., 2017) typically associated with Acheulean Large Cutting Tools (LCTs; Ambrose, 2001) and in the lower part of the range known for bifaces (e.g., minimum biface size of approximately 4.5 cm at Olorgesailie; Isaac, 1977). Given the small size (i.e., >10 cm), the morphology of the core could be characterized as a ‘protobiface’ (Leakey, 1971; Shea, 2010). This interpretation is further supported by the quasi-centripetal flaking pattern on the ventral surface (Figure 35B), following the hypothesis of Isaac et al. (1997) that protobifaces are essentially ‘elongated discoids’ (Shea, 2010). The morphology of the experimental object superficially resembles artifacts from Olduvai Gorge identified as protobifaces and LCTs (Leakey, 1971; Shea, 2010), as well as other early Acheulean core types, described as elongated cores with asymmetrical exploitation (from circa 1.3 Ma at the site of Rietputs 15 in South Africa; Leader et al., 2018). The experimental artifact also meets a further criterion used for defining ‘handaxes’, whereby “more than 2/3 of the tool’s circumference” shows evidence of “bifacial modification” (Beyin et al., 2017, p. 255).

The elongation ratio (Width/Length) of 0.74 (0.80 using the ‘box method’) is greater than in either experimental handaxes from spandrels experiments or archaeological handaxes from Olorgesailie (Isaac, 1977; Moore & Perston, 2016), while the ratio of thickness to width was

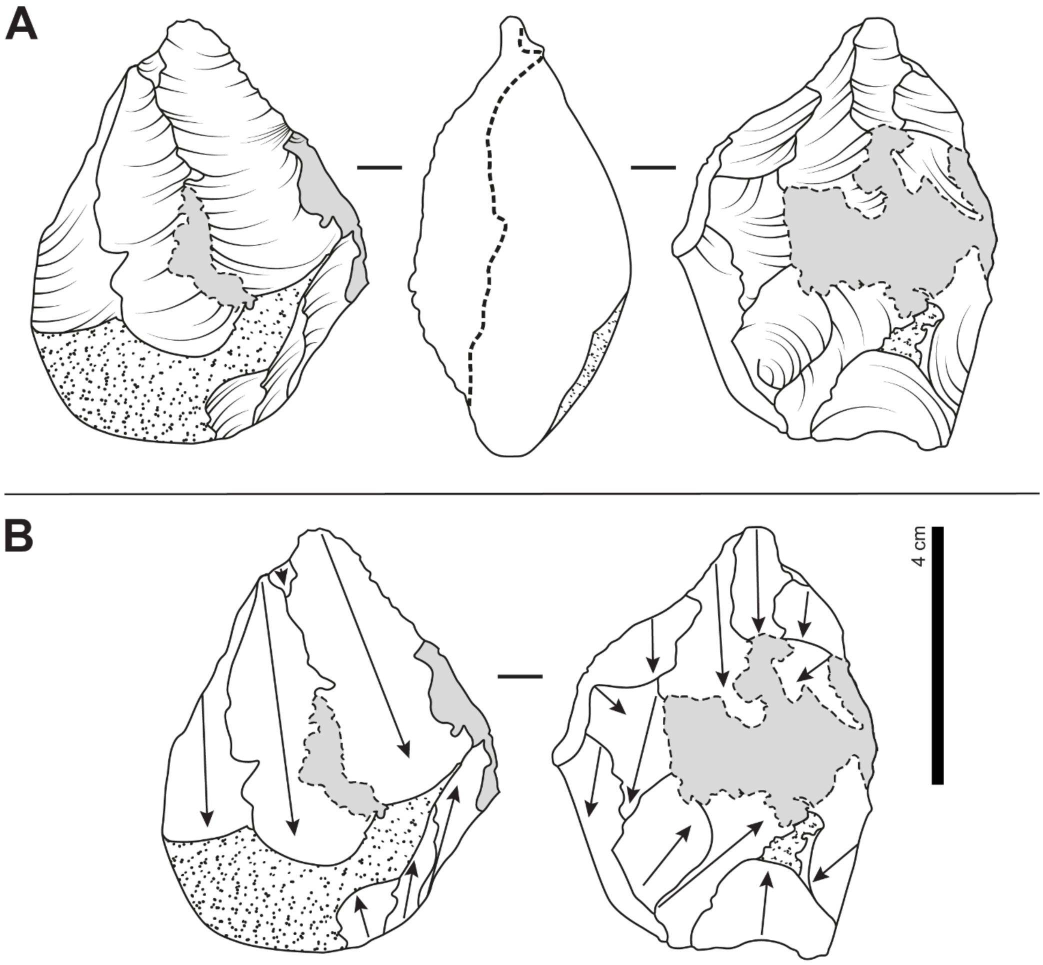


Figure 35 Illustrations of the potential protobiface (**A**), also including schematic (arrows representing directionality of flake removals) of the reduction pattern (**B**).

Table 12 Attributes of bifaces produced (presumably unintentionally) by the novice knapper (P17) and by more experienced knappers (W.D.S., main author, and D.B., see acknowledgements).

Attributes	'Unintentional Biface' (Z17.C22)	'Intentional' Bifaces	
Knapper experience	Technique-naïve (Level 2)	Experienced knappers (W.D.S., main author, and D.B., acknowledgements)	
		All replications	Excluding Method 2
Length (mm)	66.01	76.54 ± 8.19	76.55 ± 5.82
Width/Length (Elongation) (W1/L)	0.80	0.72 ± 0.09	0.70 ± 0.09
Width/Length (Elongation) (W2/L)	0.74	0.67 ± 0.10	0.66 ± 0.09
Thickness/Width (Th1/W1)	0.65	0.49 ± 0.14	0.55 ± 0.08
Thickness/Width (Th2/W2)	0.66	0.52 ± 0.15	0.58 ± 0.11

comparable to both of these samples. Again, the naïve knapper core was in the low end of the full range of Olorgesailie handaxes; spandrel experiments produced handaxes from cobbles that were much larger than the one from this naïve knapping study, but the naïve knapper core was within a standard deviation of the mean for handaxes from spandrel experiments with flake blanks (Moore & Perston, 2016).

To create a comparative sample of 'intentional' bifaces from the same glass hemisphere blanks, two knappers (W.D.S. and D.B., see acknowledgments) created nine biface replications. Seven of the bifaces were created following a similar reduction sequence as was used to generate the 'unintentional biface' of P17, albeit using exclusively freehand technique. Two bifaces were created using an alternative method innovated by D.B., which involved the flaking off of the entire proximal (flat) surface of the hemisphere. In both cases, the flake was then used as a blank, being subsequently shaped into the desired bifacial form. The unintentionally produced bifacial artifact of P17 proved fairly similar to bifaces wherefor the biface/handaxe form was imposed on purpose (Table 12).

River cobbles

Not all individuals used the provided river cobble for percussive, toolmaking activities, but many did, which is observable on numerous specimens due to abrasion and degradation from repeated hammering activities (many glass cores also show indications of these processes, though the behavior coding would indicate this can emerge both as a consequence of being used as an active element or targeted as a passive element). Several individuals (*n* = 11) produced angular fragments from river cobbles, many of which used these angular fragments to cut or sever the rope of the puzzle box (some doing so after already previously making and using cutting edge with the more easily fractured painted glass blanks, though almost individuals who modified both eventually settled on only knapping from glass). Removals from the river cobbles were not securely identified by the main experimenter as involving conchoidal fracture.

Some river cobbles were used as active elements and were also (intentionally and/or unintentionally) fractured, resulting in identifiable 'hammer-cores' (*n* = 5; Figure 36). Whereas such objects (from archaeological contexts) with percussion marks and only one or a few removals are often labelled only as 'hammers' (compare with, e.g., Delagnes & Roche,

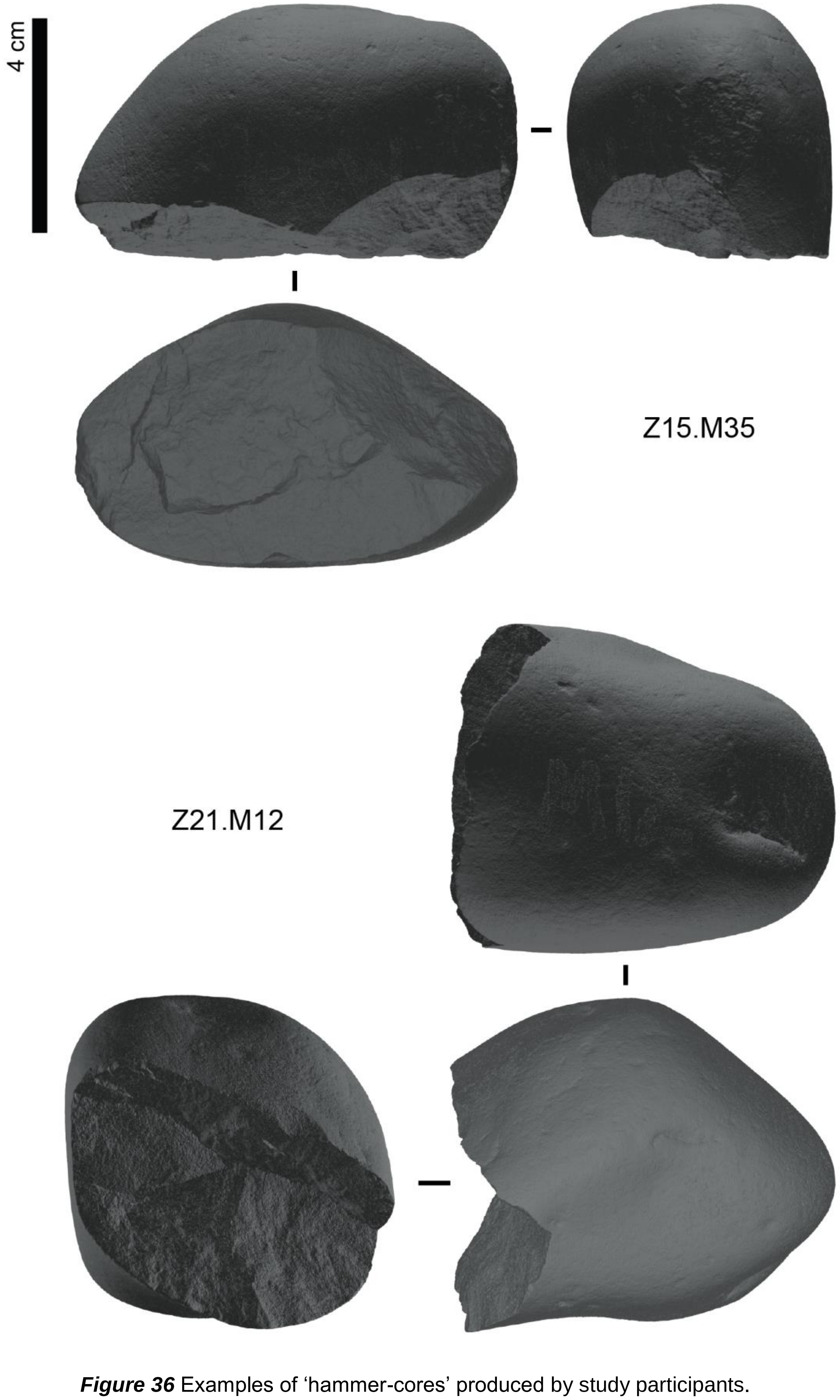


Figure 36 Examples of ‘hammer-cores’ produced by study participants.

2005), the identification of these modified river cobbles as ‘hammer-cores’ was supported by cross-validation with the behavior coding. The behavior coding indicated that at least some removals would have been intentional and the byproducts of the fracturing (i.e., angular fragments).

Principal component analysis

The results of the principal component analysis (PCA) were largely similar to those initially reported in Braun et al. (2019). The results, as expected, discriminate clearly between the technology of the Oldowan and Early Acheulean on one hand and pre-Oldowan (the site Lomekwi 3, LOM3, from West Turkana, Kenya; Harmand et al., 2015) and capuchin (SCNP; Proffitt et al., 2016) artifacts on the other.

From the output, the first principal component (PC1) explains 38.6 % of the variance in the data. Positive values are mainly related to higher percentages of cores and larger flakes. PC1 is also associated with more unifaciality, more frequent bipolar technique, a lower percentage of angular fragments, larger cores, thicker flakes, and a higher proportion of percussive objects. PC1 separates out the Lomekwi and capuchin data from the rest of the assemblages, the experimental assemblage included. The combined assemblage of the naïve knappers experiment is in the range of the Oldowan for PC1, while possessing a more negative value for PC1 than Early Acheulean sites.

The next principal component (PC2) explains 22.9% of the variance in the data. Positive values of PC1 are most strongly related to higher scar counts and higher ratios of flake scars to log (base 10) of core size. Beyond this, positive values of PC1 also predict less unifaciality, lower frequencies of bipolar cores, larger core sizes, larger flake sizes, thicker flakes, and a lower proportion of percussive objects. For PC2, the only outlier is the SCNP assemblage produced by wild capuchins with an extreme negative value, indicating more unfacial cores, more bipolar cores, smaller cores, smaller and thinner flakes, and a large proportion of percussive objects. Meanwhile, the experimental assemblage has a value for PC2 that is within range of the Acheulean contingent but more positive than for the Oldowan assemblages.

Principal component (PC3) explains 15.9 % of the variance in the data. Positive values were most strongly associated with lower ratios of flakes to core size, but also predicted higher

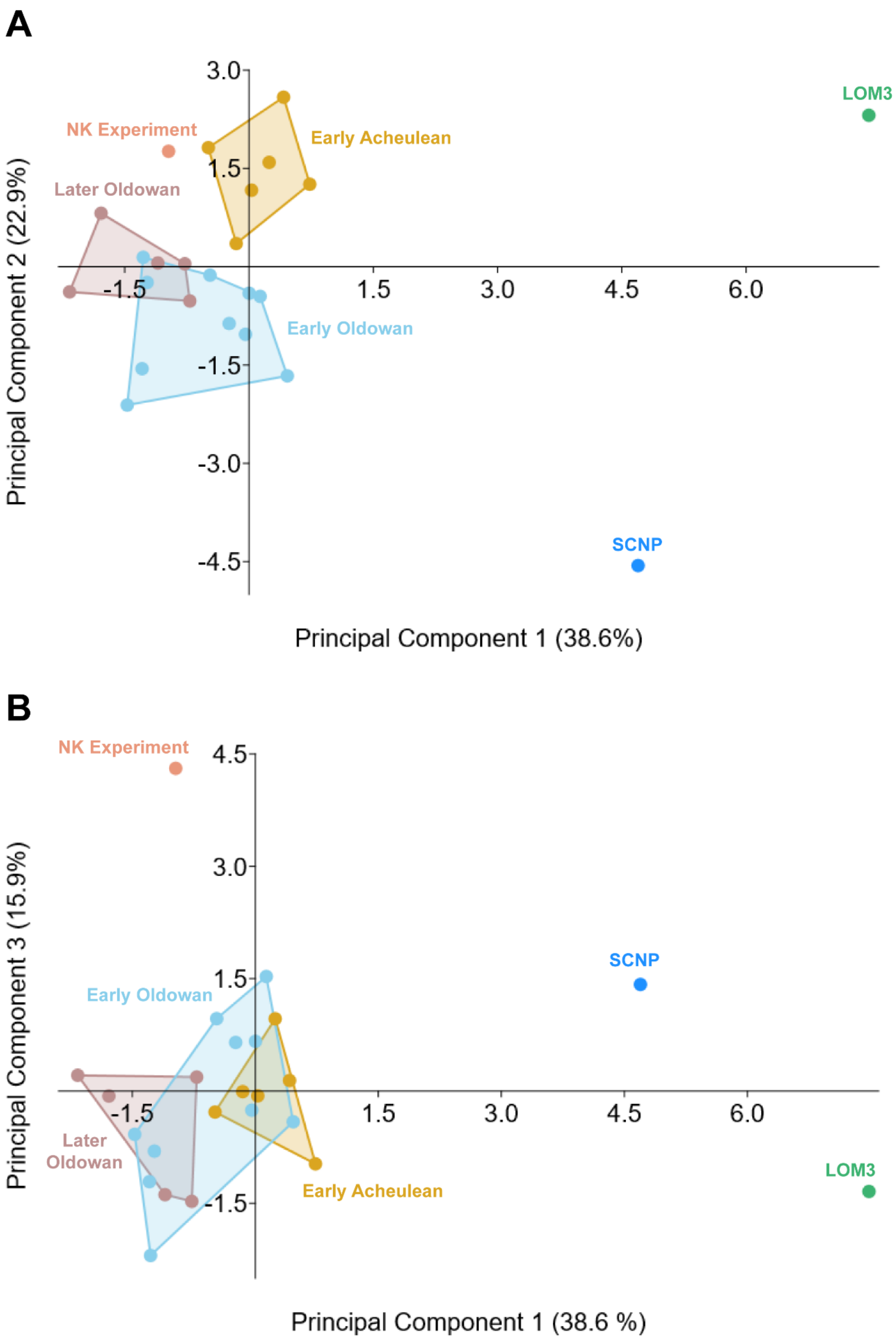


Figure 37 Plots of Principal Component 1 versus Principal Component 2 (**A**) and Principal Component 1 versus Principal Component 3 (**B**).

scar counts and more scars per log (base 10) core size. The most distinguishing characteristic of the experimental data is the high value of PC3, therefore indicating a relatively early stage of reduction and a lower proportion of flake size to core size, which might suggest more inefficient reduction, involving lots of removals of very small flakes/debitage. The input data for this PCA shows that the experimental assemblage has the highest values for both flake scar count and flake scar to core size ratio and lowest value for ratio of flake size to core size.

When plotting PC1 against PC2 (Figure 37A), the novice knapper assemblage plots most closely to Acheulean assemblages, followed by Later Oldowan and then Early Oldowan sites. When plotting PC1 against PC3 (Figure 37B), the convex hulls for the Early Oldowan, Later Oldowan, and Early sites roughly overlap, showing a generally shared technological character that distinguishes them from the assemblages produced by experimental human novices and wild capuchins, as well as the Pre-Oldowan site, Lomekwi 3. The novice knapper assemblage, with the extreme positive value of PC3, plots well away from any of the other assemblages. At the same time, SCNP and Lomekwi 3 plot closer to each other – albeit not very close at all – than to any of the other clusters or assemblages.

Besides the patterns indicated in – and ultimately not having a strong bearing on – the PCA, the experimental assemblage from novice knappers also has a higher rate of bipolar toolmaking than Oldowan and Acheulean sites (and roughly similar to Lomekwi 3).

Productivity, expediency, and efficiency of knapping techniques

According to Putt (2015), the most efficient technique for toolmaking was throwing of a hammer onto a core (indirect projectile technique), the most expedient technique was bipolar, and the most productive was freehand technique. None of the participants preferred projectile technique in either potential or confirmed toolmaking, while only one object identifiable as a flake was produced by projectile technique. Thus, the productivity, expediency, and efficiency of projectile technique could not be evaluated. Technique preferences for confirmed toolmaking that did occur among participants in our study included passive hammer, bipolar, freehand, and opportunistic knapping (no preference for any particular toolmaking technique.

In terms of productivity, there was no one technique that proved more efficient than any of the others, with a lack of statistical significance in the analyses of the ratio of flake size to flake mass (X^2 = 3.4183, df = 3, p = 0.3315, Kruskal-Wallis test), flakes produced per blank

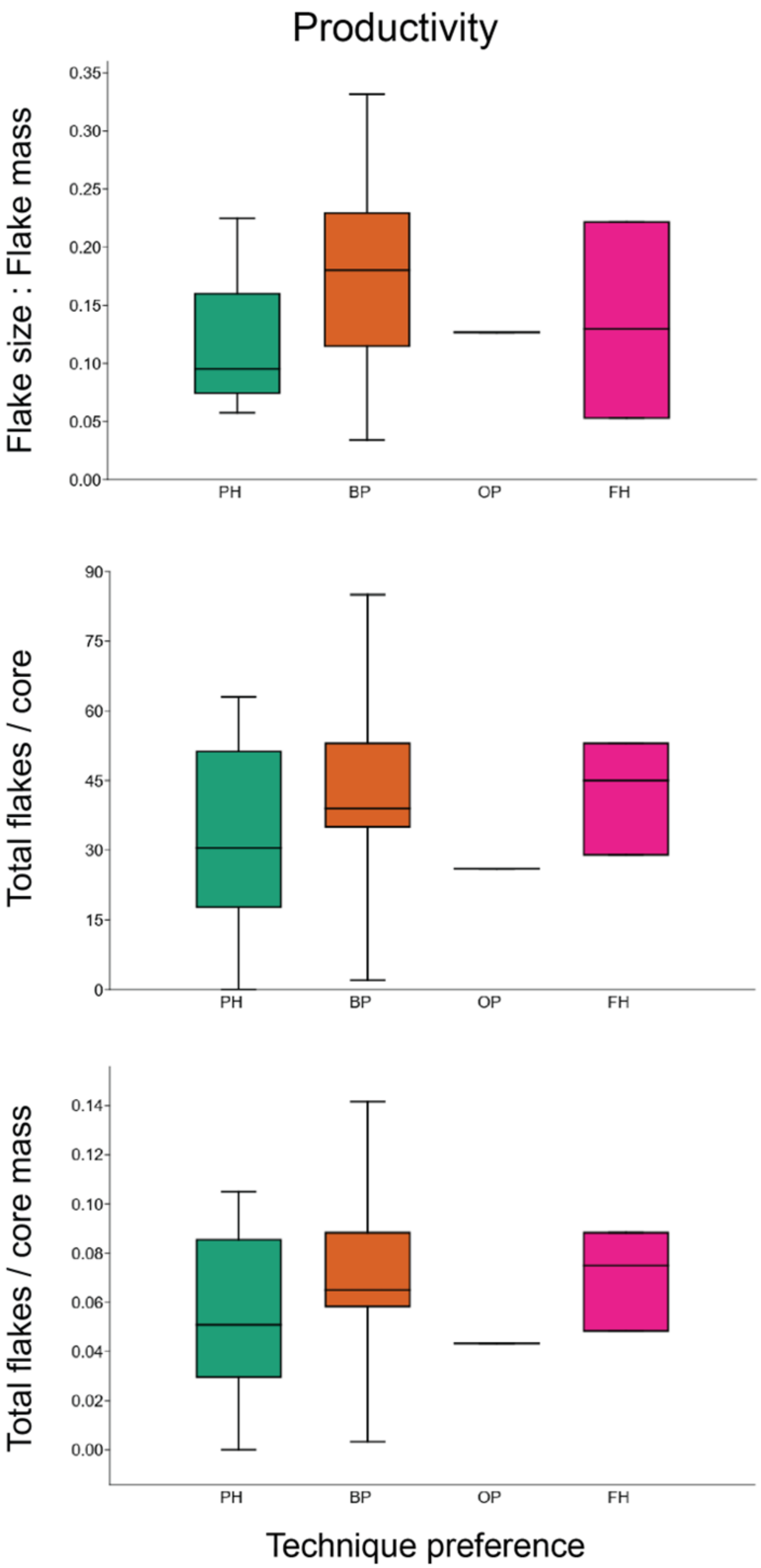


Figure 38 Box plot visualizations of participant technique preferences versus measures of productivity.

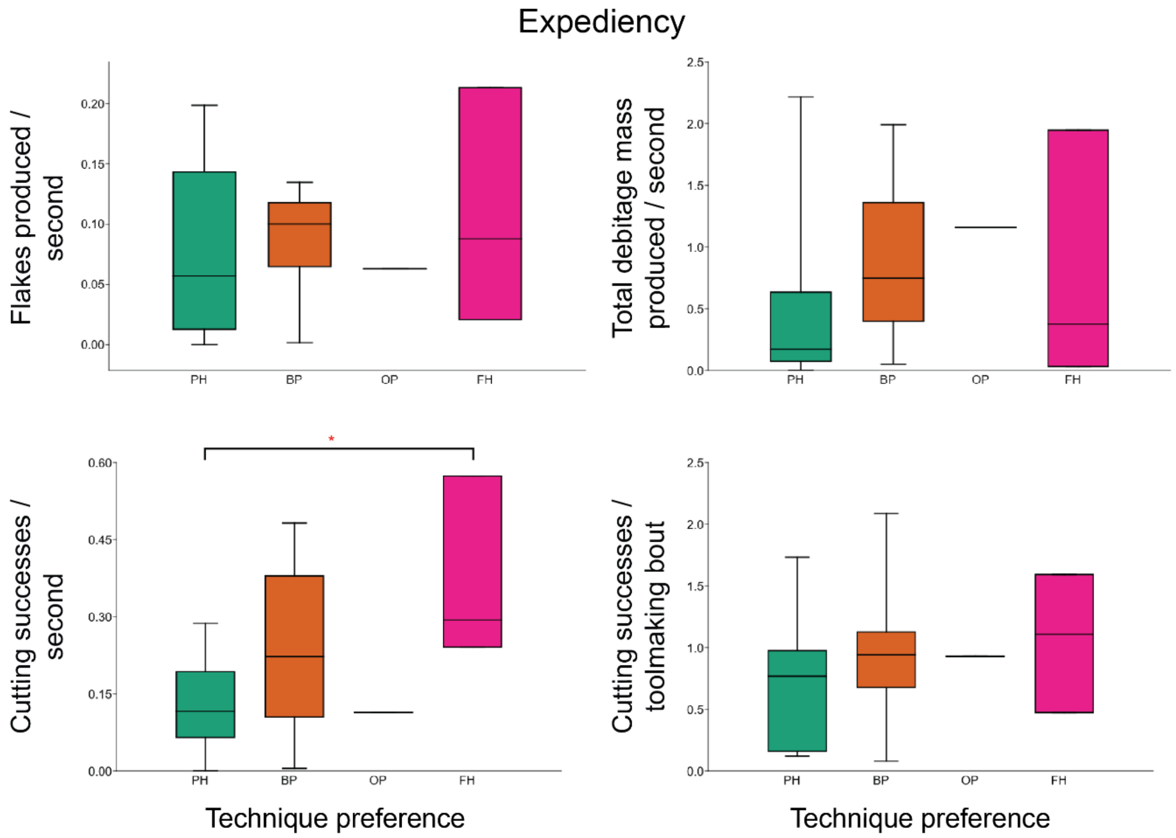


Figure 39 Box plot visualizations of participant technique preferences versus measures of expediency.

(X^2 = 3.4943, df = 3, p = 0.3215, Kruskal-Wallis test), and flakes produced per blank mass (X^2 = 3.4943, df = 3, p = 0.3215, Kruskal-Wallis test; see Figure 38).

Similarly, there was no clear advantage in terms of expediency (Figure 39) for all of the technique preferences considered. There was no statistical significance for overall comparisons of flakes produced per second (X^2 = 0.8132, df = 3, p = 0.8463, Kruskal-Wallis test), debitage mass produced per second (X^2 = 3.9131, df = 3, p = 0.2710, Kruskal-Wallis test), cutting successes per second (X^2 = 6.4353, df = 3, p = 0.0922, Kruskal-Wallis test), and cutting successes per TM bout (X^2 = 1.6656, df = 3, p = 0.6446, Kruskal-Wallis test). Pairwise comparisons, however, did reveal that there was a statistically significant difference between freehand and passive hammer techniques for total cutting successes per second (Z = -2.28192, p = 0.0225, Wilcoxon rank-sum test), with participants having a freehand technique preference being more expedient, i.e., producing more cutting successes per second.

Finally, the broad pattern for efficiency measures was much the same as for the previous two technological variable categories (Figures 40 and 41). There was no statistical significance for overall comparisons of the percentage of blank mass into flakes (X^2 = 4.7419, df = 3, p =

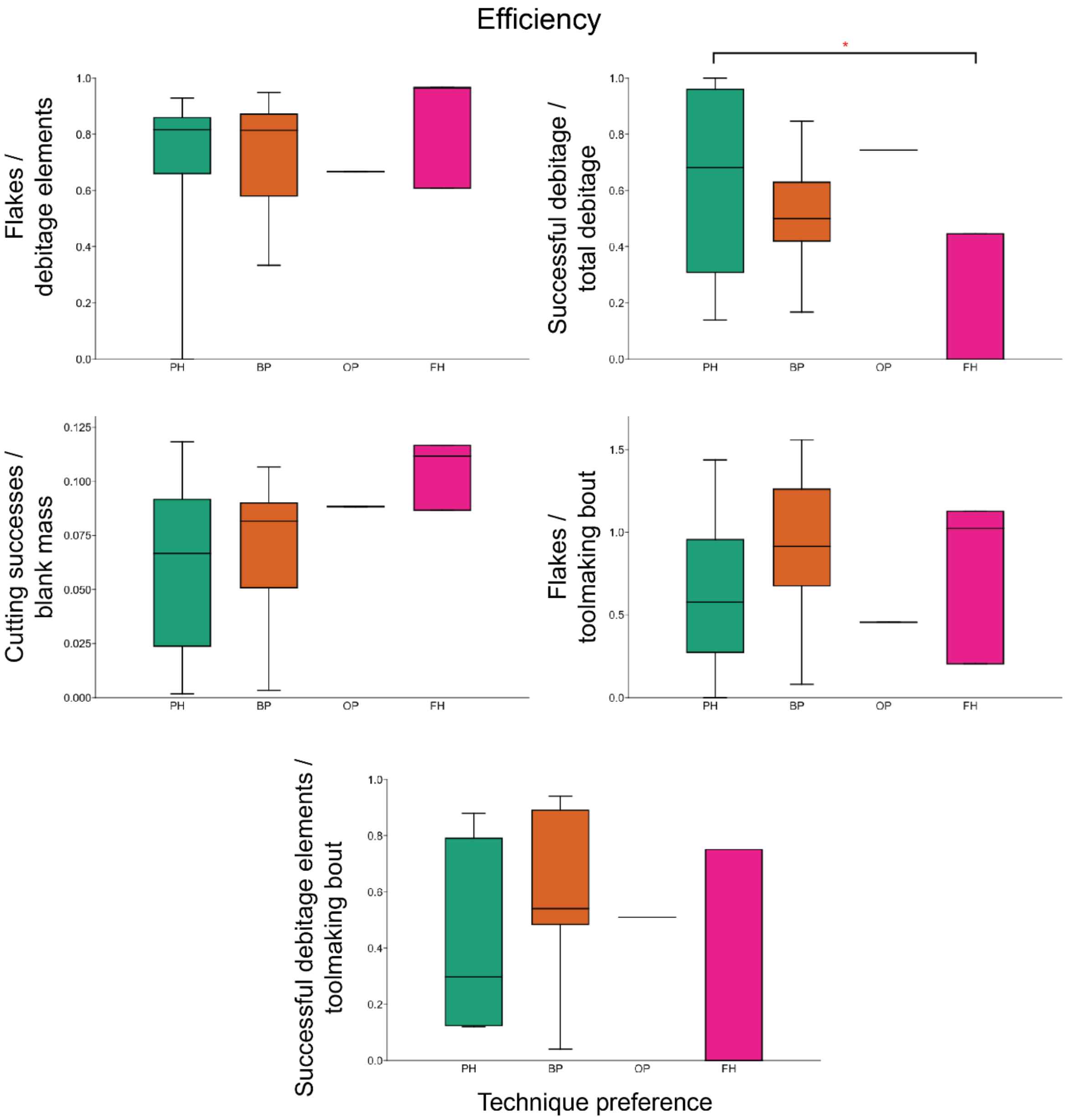


Figure 40 Box plot visualizations of participant technique preferences versus measures of efficiency.

0.1917, Kruskal-Wallis test), the percentage of blank mass into unused debitage, $X^2 = 5.1047$, df = 3, $p = 0.1643$, Kruskal-Wallis test), the percentage of blank mass into unexploited core, $X^2 = 3.7080$, df = 3, $p = 0.2948$, Kruskal-Wallis test), the proportion debitage elements that are flakes ($X^2 = 2.1075$, df = 3, $p = 0.5504$, Kruskal-Wallis test), the proportion of debitage used successfully ($X^2 = 6.3948$, df = 3, $p = 0.0939$, Kruskal-Wallis test), cutting successes per blank mass ($X^2 = 3.7356$, df = 3, $p = 0.2915$, Kruskal-Wallis test), flakes per toolmaking bout ($X^2 = 3.7564$, df = 3, $p = 0.2890$, Kruskal-Wallis test), and successful debitage per toolmaking bout ($X^2 = 5.3074$, df = 3, $p = 0.1506$, Kruskal-Wallis test). Pairwise comparisons revealed that participants with a bipolar technique preference performed significantly better

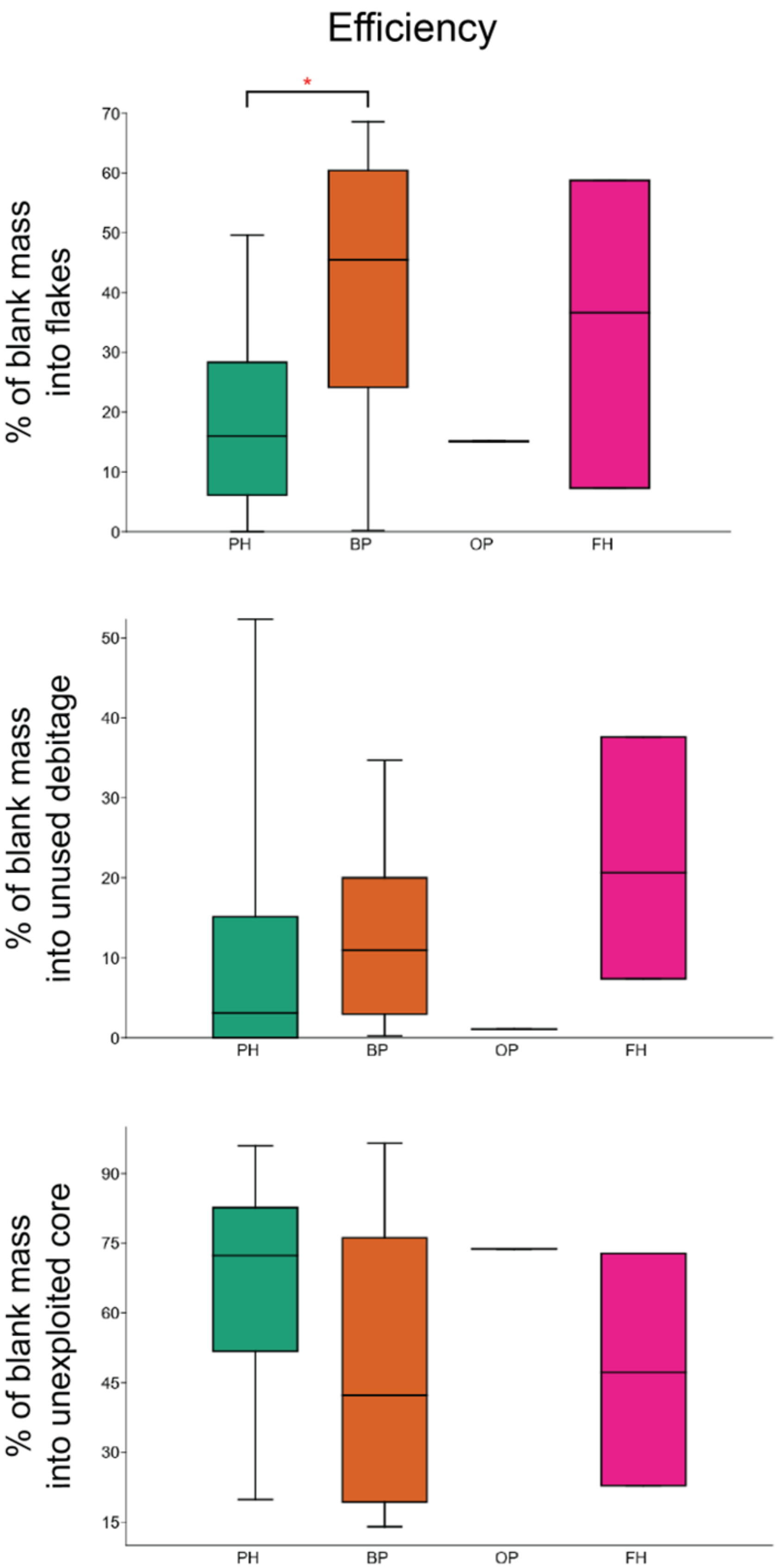


Figure 41 Box plot visualizations of participant technique preferences versus measures of efficiency.

at converting blank mass into flakes than participants with a passive hammer preference (Z = -2.07758, p = 0.0377, Wilcoxon rank-sum test) and that participants with a passive hammer preference performed significantly better with regards to the proportion of debitage used successfully as cutting tools than participants with a freehand technique preference (Z = -2.01950, p = 0.0434, Wilcoxon rank-sum test). The latter pattern may have little to do with actual toolmaking efficiency and have more to do with the fact that many participants with a freehand technique preference also happened to be preferred core-tool users and therefore did not regularly use the debitage elements they produced for cutting activities.

Attempted solutions and tool use

There were a total of 2299 attempted solutions of the puzzle box (mean by individual = 81.9, SD = 31. 4 solution attempts). Over half of all attempted solutions involved cutting actions (N = 1381, 57.0 % of solution attempts, mean = 49.3, SD = 21.4). Nearly all cutting attempts were successful (N = 1257, 95.6 % of cutting attempts).

Cutting attempts included the usage of any created or modified object as a potential cutting tool. 56 % of all cutting attempts involved the use of debitage elements (flakes or other detached pieces; labelled in the behavior coding as DB1). The remaining cutting attempts involved edges created on glass hemispheres or river cobbles. Three participants exclusively used core edges for cutting attempts (cutting attempts with DB1/total cutting attempts = 0.00; participants P23, P29, and P30). 13 participants (46.4 %) had a preference for using flake-tools, 6 participants (21.4 %) had a preference for using core-tools, and 9 participants (32.1 %) had a preference for neither flake-tools nor core-tools and were therefore considered opportunistic cutting tool users.

There was no significant difference in terms of mass for the flakes from the different tool preferences (X^2 = 4.1384, df = 2, p = 0.1263, Kruskal-Wallis test; Figure 42). Other metrics did show significant differences, including maximum dimension (F = 6.8738, df = 2, p = 0.0011, ANOVA analysis of variance). Opportunistic tool-users flakes were significantly longer than flakes made by flake-tool users (p = 0.0002, Student’s t-test). Width of flakes differed significantly (F = 6.4472, df = 2, p = 0.0016, ANOVA analysis of variance), with opportunistic tool-user’s flakes being significantly wider than flake-tool user’s flake (p = 0.0004, Student’s t-test). The difference between opportunistic tool-user’s flakes and core-tool user’s flakes was not significant, although quite close to significance (p = 0.0686,

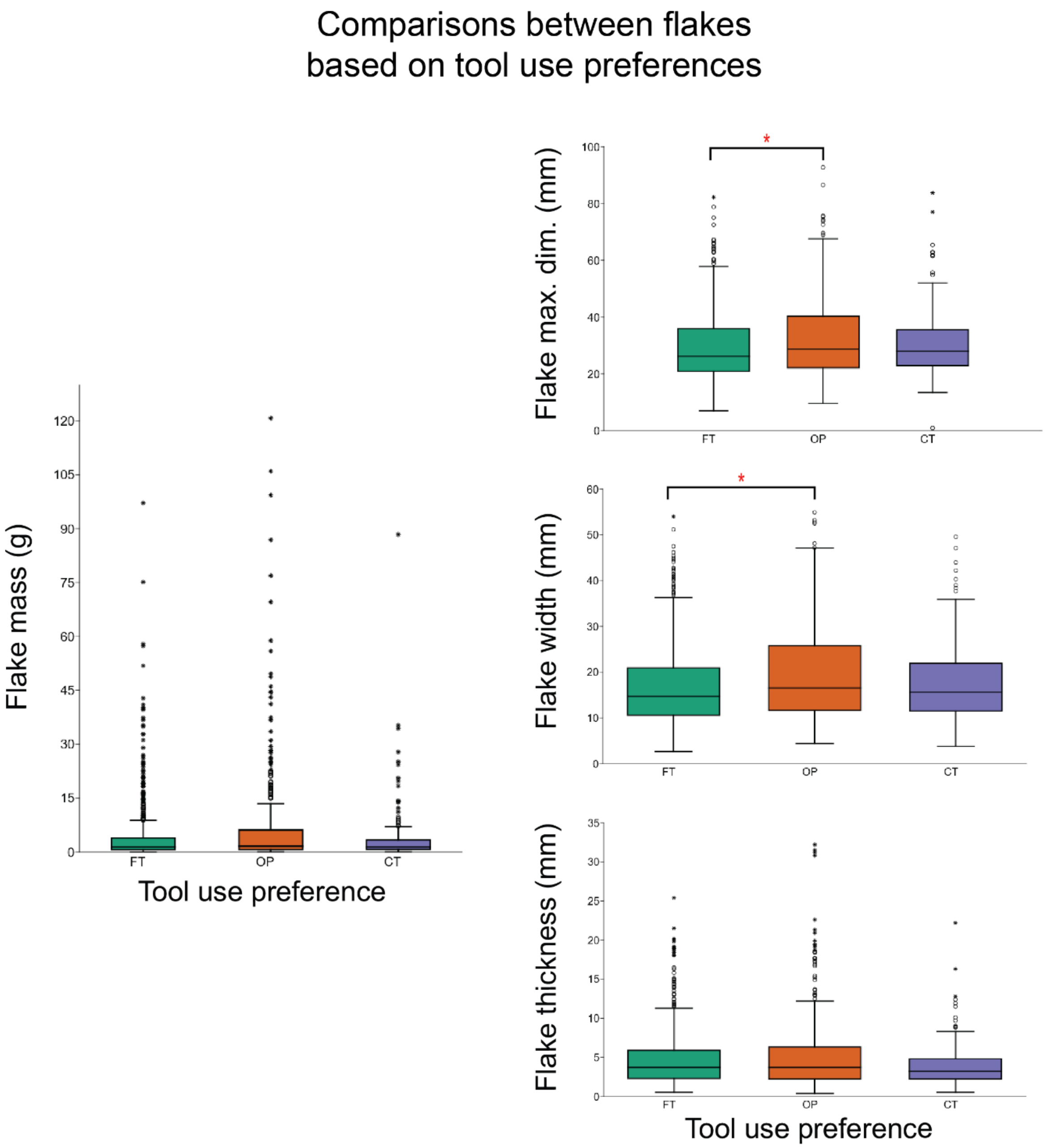


Figure 42 Comparisons between the metric attributes (mass, maximum dimension, width, and thickness) of flakes produced by study participants with different tool use preferences (flake-tool, core-tool, and opportunistic tool use).

Student's t-test), with opportunistic tool-user's flakes being again wider. Thickness of flakes differed significantly ($F = 6.7114$, df = 2, $p = 0.0013$, ANOVA's analysis of variance), with significant pairwise differences between flakes of opportunistic tool-users and core-tool users ($p = 0.0003$, Student's t-test) and between flakes of flake-tool users and core-tool users ($p = 0.0070$, Student's t-test). Both flake-tool users and opportunistic tool-users created flakes that are significantly thicker than core-tool users' flakes. There were no significant differences in terms of platform depth of flakes between different tool use preferences ($X^2 = 4.3685$, df = 2, $p = 0.1126$, Kruskal-Wallis test). Differences in terms of the final flake metric, EPA, were

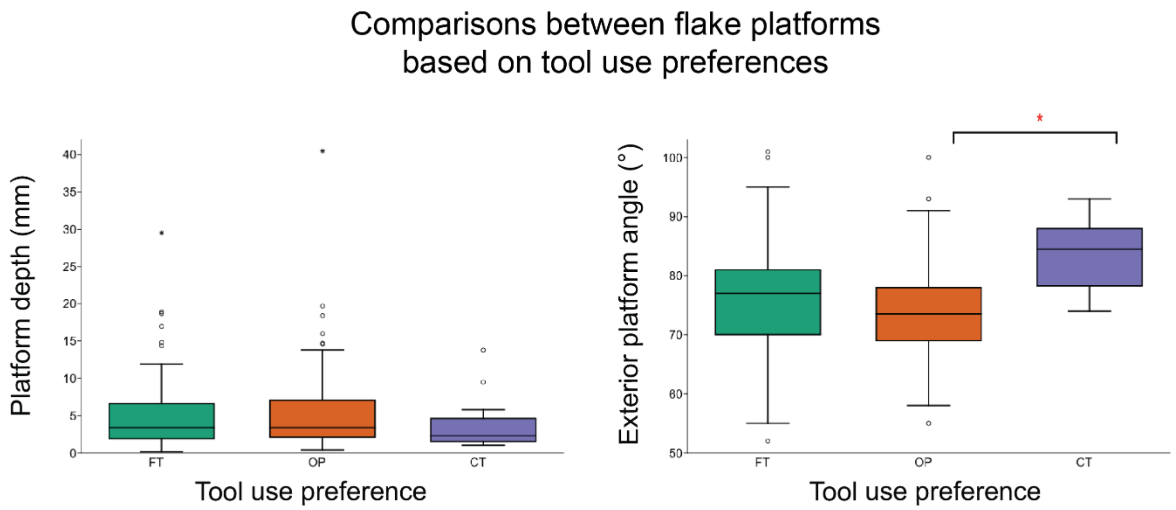


Figure 43 Comparisons between metric attributes (platform depth and EPA) of platforms of flakes produced by study participants with different tool use preferences (flake-tool, core-tool, and opportunistic tool use).

significant (*F* = 3.4854, df = 2, *p* = 0.0321, ANOVA analysis of variance, Figure 43). Core-tool users' flakes were significantly more oblique than opportunistic tool-users' flakes (*p* = 0.0106, Student's t-test) and markedly more oblique (though not statistically significant) than flake-tool users' flakes (*p* = 0.0554, Student's t-test).

Total scar count on cores differed significantly based on tool preference (X^2 = 10.3312, df = 2, *p* = 0.0057, Figure 44). Pairwise comparisons revealed that opportunistic tool users produced cores with significantly fewer flake scars than flake-tool users (*Z* = -2.15074, *p* = 0.0315, Wilcoxon rank-sum test) and core-tool users (*Z* = -2.86721, *p* = 0.0041, Wilcoxon rank-sum test). Productivity in terms of flakes produced per blank is significantly difference

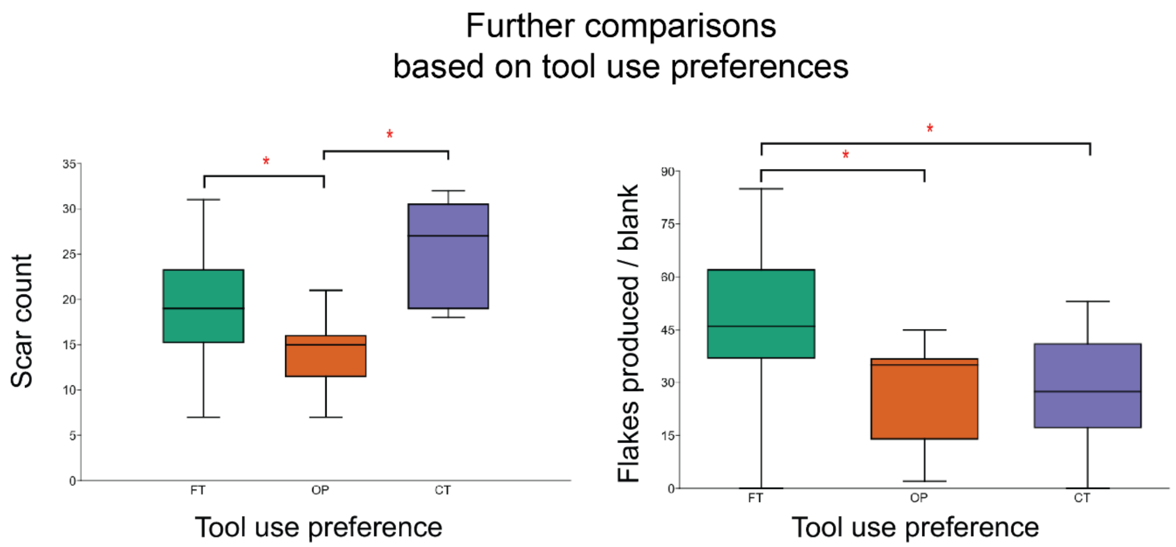


Figure 44 Comparisons of scar counts on cores and productivity in terms of flakes produced per core across tool use preference categories (flake-tool, core-tool, and opportunistic tool use).

across tool use preferences (X^2 = 7.5574, p = 0.0229, Kruskal-Wallis test). Flake-tool users produced significantly more flakes per blank than both core-tool users (Z = 1.97511, p = 0.0483, Wilcoxon rank-sum test) and opportunistic tool-users (Z = =2.41220, p = 0.0159, Wilcoxon rank-sum test).

The distributions of core types per tool use preference do not differ significantly (X^2 = 11.914, p = 0.29098, Pearson's chi-squared test). The greatest variety of core types was created by opportunistic tool-users (unifacial chopper, multifacial core, heavy-duty scraper, discoid, core-scraper, and bifacial chopper), followed by core-tool users (uifacial chopper, heavy-duty scraper, discoid, bifacial chopper), and then flake-tool users (unifacial chopper, multifacial, discoid).

Behavioral Observations

There was a wide variety of intriguing toolmaking and tool use behaviors that appeared spontaneously among our participants during testing. The following text contains just a couple noteworthy examples.

P2 created shatter and detached small pieces from the granite block by repeatedly striking it with a river cobble. She used one of the angular fragments from the granite block to cut the rope and open the puzzle box. Following this success at the task, however, she did not repeat the same toolmaking procedure. Instead, she continued striking the granite block with a river cobble, but only to produce fine silicate dust. P2 gathered 'larger' pieces and then placed them on the anvil and broke and crushed them down further into dust as well. She started covering the gloves with the silicate dust and rubbed the particles into the threads of the rope (seemingly as an abrasive agent). Attempts of this method lasted approximately 86 minutes without success, ending when the participant finally gave up on opening the box again.

P22 at one point attempted to sever the rope by hammering the river cobble against a glass flake pressed perpendicular to the rope. This hammer-and-chisel technique was ultimately unsuccessful. He also – on not just one, but a few occasions – used two flakes in combination to cut the rope, in a manner he himself described as like "scissors" (DE. "Schere").

VI. Naïve knapper study: discussion

Explanation of candidate contribution to collaborative work

This chapter is based on collaborative work. The text is partially adapted from the main text of the following publication:

Snyder, W.D., Reeves, J.S., & Tennie, C. (2022). Early knapping techniques do not necessitate cultural transmission. *Science Advances*, *8*(27), eabo2894. doi:10.1126/sciadv.abo2894

Conceptualization: C.T., W.D.S., J.S.R. Investigation: W.D.S. Visualization: W.D.S. Formal analysis: W.D.S., C.T. Writing (original draft): W.D.S. Writing (review and editing): W.D.S., C.T.

Some of the remaining content is from as-yet unpublished, in-preparation work:

Snyder, W.D., Reeves, J.S., & Tennie, C. (in prep). Untitled naïve human toolmaking and tool use paper.

According to the editors of *Science* (Special 125th Anniversary Issue, 2005), the origins of modern human-like cumulative culture of know-how count(-ed) among the top 125 open questions of our time. Various hypotheses and timelines have been proposed to answer this question, with some accounts placing said origins in a shared ancestor with apes (Schillinger et al., 2015; Shipton, 2010; Whiten, 2015, 2016, 2017; Whiten et al., 2009b), while yet others have picked out the Oldowan industry as the first verifiable sign of cultural transmission of know-how and cumulative cultural evolution (Caruana et al., 2013; Lombao et al., 2017; T. Morgan et al., 2015; Schick et al., 1999; Shipton & Nielsen, 2015; Stout et al., 2019). Experiments with human participants are believed to have shown cultural transmission of know-how was necessary for early knapping techniques in both humans and early hominins (Cataldo et al., 2018; Lombao et al., 2017; T. Morgan et al., 2015; Shipton, 2020).

The results of these experiments, however, are often contradictory and are based in circular reasoning (see Tennie et al., 2017), as the studies apply only conditions that involve cultural transmission of know-how (i.e., copying social learning opportunities, the mechanistic basis of the cumulative culture of know-how; e.g., Tennie et al., 2009, 2020; Tomasello, 1999, 2003) and therefore have lacked real control in the sense of a true baseline condition (see Tennie et al., 2016, 2017). Given these caveats and observations of stasis and technological simplicity in the early archaeological record, some researchers have alternatively suggested that cultural transmission of know-how was not present in the Oldowan (Cueva-Temprana et al., 2022; Davidson & McGrew, 2005; Horta et al., 2022; Richerson & Boyd, 2005; Tennie et al., 2016, 2017). Proponents of an alternative hypothesis consider it more likely that individual re-innovation and cultural transmission of other information types (as in primate minimal culture; Bandini & Tennie, 2017; Buskell & Tennie, in press; Neadle et al., 2017; Tennie et al., 2020) were the drivers behind knapping know-how and tool forms (Cueva-Temprana et al., 2022; Snyder et al., 2022; Tennie et al., 2016, 2017).

Contextualizing the experiment

The presented dissertation project sought to examine stone toolmaking under a novel experimental paradigm (adapted from island tests for primate technological know-how, including related to stone tools: Bandini et al., 2021c; Motes-Rodrigo et al., 2022), serving as the very first true baseline condition of human knapping skills.

Here, we recruited participants for a study on 'problem-solving abilities', rather than directly informing them about stone tools or toolmaking. There was no provision of instructions or demonstrations, nor any other information related to stone tools and knapping know-how. Participants were given motivation in the form of a baited puzzle-box, which they were told was accessible by any means they deemed reasonable, including by using the available raw materials (called 'objects' and 'things' rather than as a glass hemisphere, river cobble, and granite block). Naivety of the participants was evaluated using a post-test questionnaire, wherein their pre-study experiences involving stone tools were inquired after.

Oldowan toolmaking know-how – early knapping techniques – appeared in a cohort of technique-naïve re-innovators (n = 22). Conceptual knowledge of stone tools (e.g., Apel, 2001; Bamforth & Finlay, 2008; Harlacker, 2003; Pargeter et al., 2020) proved to be completely irrelevant for the reproduction of knapping techniques and the use of cutting tools. Instead, all that is required is valid raw materials and sufficient and/or suitable motivation. The capacity for recognizing raw material affordances being broadly associated with extinct toolmaking hominins (e.g., Delagnes & Roches, 2005; Goldman-Neumann & Hovers, 2011; Lombao et al., 2017; Stout & Semaw, 2006; Stout et al., 2010; comment by Shipton on Tennie et al., 2017). Study participants were capable of spontaneously inducing conchoidal fracture, achieved via identification of appropriate platforms and angle for creating flakes.

All four early knapping techniques – passive hammer, bipolar, freehand, and projectile – were shown to be individually reinnovatible – i.e., capable of being developed in the total absence of opportunities for the cultural transmission of know-how (contra, e.g., Apel, 2001; Lombao et al., 2017; T. Morgan et al., 2015; Shipton, 2020). Three of the four early knapping techniques (passive hammer, bipolar, and freehand) were used by the two totally naïve participants, while the fourth technique, projectile, appeared in the larger sample of technique-naïve individuals.

The artifactual outcomes of the participants' behaviors were also reminiscent of those from the Oldowan, following reduction sequences and generating artifact types, which are known from the archaeological record for Mode 1 technologies (de la Torre, 2011; Leakey, 1971; Moore & Perston, 2016; Toth, 1985).

External validity of the experimental conditions and raw materials

For the results of experimental research to hold any epistemological worth, it should ideally be possible to generalize said outcomes to the real-world present and past (external validity; see for example, Eren et al., 2016; Lin et al., 2018; Pargeter et al., 2019).

A major caveat for any knapping study is that living humans are distinct from the original – extinct – toolmakers in terms of biology, behavior, cognition, *and* culture and are therefore imperfect models for building inferences about the processes – cognitive and otherwise – involved in prehistoric toolmaking (e.g., Bandini et al., 2022; Davidson, 2017; Killin & Pain, 2022; Putt, 2015; Putt et al., 2017; Schick & Toth, 1994; Stout et al., 2009; Schick & Toth, 1993). Although living modern humans (and other primates) are imperfect stand-ins for extinct hominins (e.g., Bandini et al., 2022; Hovers, 2012; Killin & Pain, 2022; Putt et al., 2017; Stout & Semaw, 2006), experimental archaeology still provides insights that cannot be gleaned from studying artifacts alone – as such, experimental archaeology is a necessary means of constructing middle-range theories for how these artifacts were created in the past (Binford, 1972; Braun et al., 2006; Killin & Pain, 2022; Stout & Khreisheh, 2015). So long as the reality of the differences between living humans and extinct hominin progenitors are kept in mind throughout the research process (i.e., in experimental design, analysis, and interpretation of results; this is too often given too little conscious consideration), humans are appropriate replacements for extinct toolmakers. Being the phylogenetically closest living relatives of all other toolmaking hominins, living humans are also the best available option.

In our study, the participants were WEIRD, which in most cases might be regarded as problematic (see Bandini et al., 2022; Henrich et al., 2010; Killin & Pain, 2022). We consider it to be ideal – given the available choices – for our proof-of-principle, as it allows for more control over past experiences and pre-existing know-how within our participants that might confound the results of the experiments (i.e., people growing up in WEIRD societies, do so in physical and cultural isolation from 'indigenous' stone toolmaking contexts, cf. Stout, 2002).

For this study, we utilize what we consider to be an ecologically-grounded approach. In the experiments, the baited puzzle box acted as a simulation of extractive foraging (with money as a replacement for food-based nutrients). The expression of toolmaking behavior is therefore an emergent property of the experimental setup, being motivated extrinsically by the prospect of a reward (as in naïve primates studies: Bandini et al., 2021c; Motes-Rodrigo et al., 2022; Westergaard & Suomi, 1994a, 1995b) rather than social-oriented motivations (e.g., like toolmaking apprenticeships in some living societies: cf., Bamforth & Finlay, 2008; Stout, 2002) or intrinsic motivation (e.g., personal interest in prehistory or toolmaking: cf., Pargeter et al., 2019; Stout & Hecht, 2017). Despite prior remarks on the matter (cf. Nonaka et al., 2010; comment by Davidson on Tennie et al., 2017), knapping experiments have generally treated toolmaking as a behavior totally divorced from its intended purpose (i.e., as is assumed from mainstream interpretations of Oldowan and Mode 1 toolmaking in the archaeological record: e.g., Ambrose, 2001; Hayden, 2015; Toth, 1985; Wynn & McGrew, 1989). Only a few rare exceptions have provided human participants with materials upon which the cutting utility of created tools can be tested (Bisson, 2001; Stout & Semaw, 2006; Stout et al., 2009), but even here, the motivation is reversed, because the participants were asked to flake or perform knapping activities firsthand instead of flaking being an intermediate step in a chain of operations leading to foraging tool use.

As far as the selection of raw materials, there is an inescapable tradeoff between raw materials that are completely external valid (i.e., using the same exact rock types and forces that were available to and use by Oldowan hominins) and experimental control, which allows for reliability (i.e., raw materials that are consistent in internal properties, shape, and size, as well as accessible in sufficient quantities for knapping tests) and higher precision in the measurement of outcomes (i.e., higher internal validity). In our preparations, we prioritized the latter, ordering hemispherical paperweights from glass wholesalers, which were then covered in grey spray-paint.

Our participants were provided with these hemispherical glass blanks, which were tested beforehand for their suitable as blanks for basic flaking procedures and early knapping techniques. Though soda glass in the form we used was not available to any hominins during the Paleolithic, it was a suitable raw material for our experiments, due to the fact that glass conchoidally fractures much in the same way as other knappable materials (Dogandžić et al., 2020) and volcanic glass, or obsidian, was available at least at some points in time during the Oldowan (Piperno et al., 2009). The hemispherical shape of the blanks was also judged valid, since similar shapes (commonly referred to as split-cobbles) have been associated with Oldowan sites (Toth, 1985). This choice of blank form was validated by the production of all core outcomes that have been recognized as typical products from initially hemispherical blanks (Toth, 1985). Although the material (glass) and shape of the blanks (with available knapping platforms) could lead to interpretation that the blanks are exceptionally flakeable, it should be noted that Oldowan hominins within and across sites were exposed to raw materials that were also variably suitable in terms of material flakeability and flaking platform accessibility (cf. Delagnes & Roche, 2005; Gabrić et al., 2021; Goldman-Neumann & Hovers, 2011). While blank form certainly has the potential to drive the behavior and outcomes of knapping individuals (perhaps centripetal flaking sequences were an obvious result when using hemispherical blanks), it has also has the potential to constrain the variability in strategies and end-products. We would predict that, when provided different raw material types and forms, naïve human participants would still be able to re-innovate knapping techniques (perhaps not with the same exact frequency observed here) and those Oldowan artifact types and reduction strategies not identified from the knapping outcomes of our participants (de la Torre, 2011; Leakey, 1971; Toth, 1985).

Regardless of the caveats and considerations described above, we still consider our approach – as the first true application of a baseline methodology (Tennie et al., 2016, 2017) for human knapping abilities – to be sufficiently valid as to have consequences for our understanding of past hominin toolmaking and cognition.

Comparisons to previous studies

Previous experimental studies in cognitive archaeology have concluded that copying social learning was necessary for toolmaking know-how acquisition and skill development, even in the earliest Oldowan (Cataldo et al., 2018; Lombao et al., 2017; T. Morgan et al., 2015; Stout et al., 2019). Here, circular reasoning may be at play (Tennie et al., 2017): by assuming that certain behaviors and cognitive mechanisms were already required to produce a particular artifactual outcome (e.g., T. Morgan et al., 2015; Shipton, 2020; Stout et al., 2019), finding those artifacts will then only prove that assumption true. Past experiments have consistently provided conditions that are known to be favorable for learning by modern humans (cf. Pargeter et al., 2019, 2020, 2021), so improved performances and more 'efficient' learning when given opportunities for cultural transmission of know-how may only demonstrate how WEIRD modern humans learn best to adopt technical, outcome-oriented know-how (cf. Cataldo et al., 2018; Lombao et al., 2017; T. Morgan et al., 2015; Putt et al., 2014).

The time participants in our study needed to discover the fracture properties of the material(s) by applying suitable techniques or knapping gestures was compatible in its shortness with results of other studies with different learning conditions (similarly requiring only a few minutes: T. Morgan et al., 2015; or otherwise quite quick: Bargallo & Mosquera, 2014;

Bargallo et al., 2018; Stout, 2006; Stout & Semaw, 2006), but clearly disproves conclusions from anecdotal observations and other studies that have suggested that "totally inexperienced knappers are even unable to strike the blank at the right angle to obtain conchoidal fracture" (Lombao et al., 2017, p. 8; see also Shipton, 2020; Sterelny & Hiscock, in press; Schick & Toth, 1993). Opportunities for copying social learning and cultural transmission of know-how in other studies (Cataldo et al., 2018; Lombao et al., 2017; T. Morgan et al., 2015) thereby had an accelerative effect on the acquisition of know-how related to early knapping techniques. Ultimately, the elongated learning time of four hours that we provided our participants turned out to be more than was required for individual re-innovation, as the initial toolmaking bouts tended to occur in less than tens of minutes. Expedient early progress in skill acquisition (Apel, 2001; Geribàs et al., 2012; T. Morgan et al., 2015; Putt et al., 2014, 2019; Stout, 2006; Stout & Semaw, 2006; Stout et al., 2009; Pargeter et al., 2021) followed by a plateau(s) (unfortunately, longitudinal experimentation could not be carried out here) seems to be fairly typical across novices learning to make Lower Paleolithic artifacts regardless of the learning conditions applied (Pargeter et al., 2021; Putt et al., 2019; Stout & Khreisheh, 2015; Stout & Semaw, 2006; Stout et al., 2009).

The provision of opportunities for the cultural transmission of knapping related know-how would have influenced the general expression of knapping know-how, including the variability in the knapping techniques implemented by the participants in those studies. Except in those studies where other techniques (such as passive hammer or bipolar) were specifically targeted by the testing program (Duke & Pargeter, 2015; B. Morgan et al., 2015; Putt, 2015), the standard technique used in knapping studies has been the one also most often implicated in the behavior of Oldowan and later hominins, freehand technique. This is despite the fact that other techniques are known to have been used during the Oldowan and also in later periods (de la Torre, 2019; Putt, 2015; Toth & Schick, 2011, 2018). Additionally, there is the credible possibility that our interpretation of the techniques from archaeological evidence may be based on incomplete reference samples, thus resulting in an overestimation of freehand technique with a complementary underestimation of other toolmaking behaviors and techniques (cf. Byrne et al., 2016; Duke & Pargeter, 2015; Hiscock, 2015; Pargeter & Eren, 2017; see also below).

As a consequence of explicitly prescribing one technique for the participants to adopt, the participants in other studies likely were then restricted from more flexible expression of toolmaking know-how, with freedom for experimentation and variable adaptation to task constraints being an ostensibly key component to (knapping) skill acquisition (see relevant work on interindividual variation in knapping gestures, including between learners and their teachers: Rein et al., 2014). Nonetheless, there are some examples of anecdotes and more controlled studies where novices re-innovated bipolar technique despite only ever seeing freehand knapping being applied (Ferguson, 2003; Geribàs et al., 2010; Pargeter et al., 2021; Sternke & Sørensen, 2007). The lack of differences in performance between technique-naïve and non-technique-naïve participants in the present study (Snyder et al., 2022; contra observed differences in Putt, 2015), combined with the aforementioned evidence, indicates the most important component of knapping skill acquisition is not so much direct copying of a demonstrator's actions and outcomes, but rather the process by which one individually adapts to specific task constraints, raw material opportunities, and their own unique physical and cognitive limitations and affordances (Bril et al., 2010; Nonaka et al., 2010; Rein et al., 2014). By this account, the conscious focus on freehand technique (or some other technique or gesture) would have only increased the frequencies of the expected know-how in the study groups (mostly) at the expense of other techniques that might have played a larger developmental role than is currently ascribed.

In fact, this creates the possibility that poor performance of some novices in comparison to other novices, modern expert knappers, and purported hominin expert knappers (e.g., Stout & Semaw, 2006; Stout et al., 2009) was the result of them being restricted in their ability to apply techniques, gestures, and methods that are perhaps easier to acquire and apply for making cutting tools by novices who have not yet developed more fine-tuned motor skill via longer-term practice opportunities (Bril et al., 2010; Nonaka et al., 2010; Pargeter et al., 2019, 2020; Stout et al., 2009). This possibility is again supported by re-innovation of bipolar technique by novices exposed only to freehand technique, as bipolar technique is suggested to have been more expedient (e.g., Gurtov & Eren, 2014; Putt, 2015) and also easier and less skill-intensive as a toolmaking method (cf. Gurtov & Eren, 2014; Hiscock, 2015; Horta et al., 2022; B. Morgan et al., 2015; Putt, 2015). Furthermore, our study participants were freed to switch between techniques as cores changed in size and shape throughout the reduction process, allowing them to apply techniques that are more suitable when cores are heavily reduced and therefore lacking in workable knapping platforms with more accessible angles (B. Morgan et al., 2015). The lack of considerable differences between productivity, expediency, and efficiency among participants categorized by technique preferences might indicate also that individuals select strategies that maximize their initial sensorimotor capacities. With enough time, they might adopt different gestures and techniques as their faculties become more attuned to the task constraints (Bril et al., 2010; Nonaka et al., 2010; Rein et al., 2014).

Pre-existing social information might also impede – especially, WEIRD – individuals from spontaneously expressing knapping-related know-how (cf. Bandini et al., 2022; Henrich et al., 2010). This is evidenced, for example, by participants who verbally inquired whether it was permissible to break objects and especially by those who reacted apologetically when they broke something for the first time during testing (similarly, two participants from another study were cautious in their flaking behaviors due to fear of self-injury: Gabrić et al., 2021). Naturally, pre-existing know-how also has the potential to induce subsequent know-how (e.g., previous knowledge about the breakability of rocks or the utility of a sharp object to cut a rope). This is, however, a weak critique of our study. First, recognition that sharp objects exist and can be used for cutting represents causal inference and innovation that can be reached by individual learning alone and does not itself require cultural transmission to occur. Second, participants in our study – despite having this likely, yet assumed, pre-existing knowledge – rarely sought production and use of a cutting tool as their initial solution of the task (n = 3), instead requiring an average of 17.9 non-toolmaking puzzle box solution attempts prior to the first occasion of confirmed toolmaking. Thus, the island test experiments were opaque enough be secure about the validity of the re-innovations, but not too opaque that no toolmaking occurred at all (contra e.g., Shipton, 2020). Overall, a channeling of the participants towards a perceived need to cut or make a cutting tool still is better regarded as motivation for the re-innovation techniques, as it does not directly transmit any of the know-how for early knapping techniques.

Cognitive archaeologists have placed considerable emphasis on the role of knapping errors for the identification of novice knapping products in prehistory (e.g., Schick, 1994; Shelley, 1990; Stout et al., 2009). In line with that reasoning, experimental studies have found that novice products are distinguishable from modern expert products and real Oldowan artifacts in terms of both knapping errors and other metric and qualitative attributes (e.g., Pargeter et al., 2021; Stout & Semaw, 2006; Stout et al., 2009; note, however, the consequential observations of Proffitt et al., 2021b). Similar markers of so-called inexperience to those documented in earlier research (Bril et al., 2010; Nonaka et al., 2010; Shelley, 1990; Sternke & Sorensen, 2007; Stout & Semaw, 2006; Stout et al., 2009), such as step and hinge fractures, indications of battering, and relatively non-invasive flaking of the cores, were found on the knapping products of our participants.

From these observations and patterns (and from comparisons with primate toolmaking, see below), the most parsimonious explanation is that early knapping techniques do not necessitate cultural transmission of know-how to be expressed by naïve modern humans (Tennie et al., 2017). Growing up in culture-enriched environments where these mechanisms and strategies for copying social learning are ubiquitous enables modern humans to more quickly and effectively acquire know-how and adopt skills that are otherwise within the reach of individual re-innovation (cf. Bandini et al., 2022; Neldner et al., 2020).

Comparisons to extant primate data

Early attempts to elicit the expression of Oldowan-like toolmaking in apes involved training of the apes (an orangutan and bonobos) by human experimenters (Wright, 1972; Toth et al., 1993). Despite the continuous training with demonstrations and other social input from the experimenters, the bonobos Kanzi and Panbonisha were able to produce only ‘low-quality’ artifacts that ultimately were well outside the artifactual variation known from the Oldowan (Stout et al., 2009; Toth et al., 1993, 2006). The validity of these studies is also fundamentally questionable, because the tested individuals were all highly enculturated (i.e., raised in captivity by humans) and therefore non-representative of the cognitive and behavioral capacities of their wild counterparts (Bandini et al., 2020, 2022; Henrich & Tennie, 2017; Tennie et al., 2020). More recent tests of naïve, unenculturated primates have shown mixed results. Chimpanzees, using a paradigm similar to the one for naïve humans (Snyder et al., 2022), were not capable of either re-innovating any early knapping techniques or utilization of cutting tools (Bandini et al., 2021c). In separated test conditions, orangutans were shown capable of passive hammer technique to make flakes and cutting tool use (Motes-Rodrigo et al., 2022). Finally, capuchin monkeys have demonstrated the highest propensity (among non-human primates) for knapping know-how re-innovation. Naïve capuchin monkeys independently and spontaneously created flakes using variants of passive hammer, bipolar, freehand, and projectile techniques, while also being able to use flakes as cutting tools (Westergaard & Suomi, 1994a, 1995a).

Wild primates have been observed engaging in tool activities that produce valid conchoidal flakes (Luncz et al., 2022a; Proffitt et al., 2016, 2021a). These, however, do not involve the intentional production of cutting tools, as the flakes are merely byproducts of behaviors with other purposes (Luncz et al., 2022a; Proffitt et al., 2016). The results of the naïve knapping experiments provide an intriguing parallel to the circumstances of flaking by wild primates, given the frequent occurrence of simultaneous toolmaking and use, especially as the initial condition of toolmaking. The clear difference is that once the properties of the material were discovered via an initial fracture event, human participants were nearly guaranteed to follow-up with repeated instances of making and using cutting tools, while the same cannot be said for the primates. Also, there is a peculiar similarity between the silicate dust production by P2 (presumably with the intent to use the dust as an abrasive element) and the behavior of wild capuchins that seemingly intend to produce dust as well (and which they then ingest, for an unclear purpose; Proffitt et al., 2016).

When compared to their primate equivalents, WEIRD human participants are exceptionally more reliable in their abilities to re-innovate knapping techniques and cutting tool use. This clear difference in re-innovation frequency and variability can potentially be explained by several different factors, including the enculturation of living modern humans increasing the likelihood of exploration and innovation in prescribed testing conditions or the biological and cognitive differences between the disparate taxa, e.g., the relationship between innovation and encephalization (Bandini et al., 2022; Henrich et al., 2010; Neldner et al., 2020; Reader & Laland, 2002). It is possible, for instance, that humans and their hominin forebearers have improved capacities for making causal inferences that are relevant to expression of toolmaking know-how (de Beaune, 2004; Placi, 2022). Pre-existing know-how might also restrict potential for re-innovating knapping, as fractured tools are typically abandoned for being less effective percussors (e.g., Carvalho et al., 2008). Re-innovation of early knapping techniques and knapping outcomes appears to be within the reach of at least some primate species, but likely only exist at the very fringe of their respective ZLSs, being difficult and/or unlikely to develop individually (cf. Neadle et al., 2020; Tennie et al., 2020).

Even if knapping is relatively infrequent in primates (varying with both taxon and context) and typically not accompanied by subsequent cutting tool use, primate data provides crucial evidence that both direct *and* indirect cultural transmission of know-how are unnecessary for development of the target know-how.

Comparisons to the archaeological record

Interpretation of behavior and cognition from the early archaeological record is complicated by the very nature of the record itself. Oldowan sites, especially, tend to have formed as palimpsests and are often time-averaged accumulations of hundreds and thousands of years of hominin activities (cf. Dominguez-Rodrigo, 2009; Duke & Pargeter, 2015; Pargeter et al., 2019; Perreault, 2019). The Oldowan *sensu latu* spans for over a million years and includes sites throughout Africa and Eurasia, where hominins of different taxa utilized a wide assortment of raw materials to create sharp-edged tools. While actualistic experiments have mainly concentrated on the reduction of cores by single agents in short, restricted time periods of minutes or hours (barring a few meager exceptions where more than one individual was part of a knapping sequence, e.g., Ferguson, 2003; Shipton & Clarkson, 2015), Oldowan artifacts may be the remnants of the accumulated knapping actions of any number of individuals of varied skilled levels (thus masking the actions of less skilled knappers and creating an illusion of proficiency and skill as a general rule for early Oldowan toolmakers; see Ferguson, 2003; Shelley, 1990; Stout & Semaw, 2006; Torres & Preysler, 2020) and perhaps even from different species of hominins (compare to raw material and tool 'sharing' in primates, e.g., Fragaszy et al., 2013; Reeves et al., 2021). Further problems with interpreting behavioral and cognitive processes underlying prehistoric stone tools arise from the interplay between raw material and skill, with current evidence indicating this to be a dynamic and quite confounding relationship (see Ferguson, 2003; Flenniken, 1984; Harlacker, 2003; Proffitt et al., 2021a; Stout & Semaw, 2006; Stout et al., 2009; Toth et al., 2006).

As such, direct comparisons between artifacts from experimental conditions and any specific or even all actual Oldowan sites should be considered with these caveats in mind. Following this line of thought, the (external) validity of our study's results for theoretical implications on the Oldowan are judged on four criteria (always remaining aware of the specific conditions of the experiment, e.g., raw materials included, and how they may influence the outcomes):

1) Are artifacts from naïve participants under island test conditions similar to those from participants in other studies who had been given copying opportunities? And do the artifacts made by our participants and by participants in other studies differ in a similar fashion to real Oldowan artifacts?

2) Do the detached pieces meet the minimum standard of external validity: flakes generated by conchoidal fracture?
3) Do the participants use the same early knapping techniques that are assumed to have also been used during the Oldowan?
4) Is the technological character of the cores and the whole assemblages reasonably similar to that from the Oldowan?

The answer to all these questions is – unequivocally – yes.

Earlier experimental studies have shown a marked distinction between the artifacts made by human novices on one side and the artifacts made by human experts and from the archaeological record on the other (Pargeter et al., 2019, 2021; Stout & Semaw, 2006; Stout et al., 2009). This is true even when novices have been provided learning conditions considered ideal for WEIRD individuals, with both ‘ample’ time (approximately two hours) and copying opportunities (Pargeter et al., 2021). Now, some Oldowan sites do appear to have been the consequence of limited skill on the part of the knappers or instead of shoddy raw materials (see Delagnes & Roche, 2005; Stout et al., 2009). Indeed, the affordances of available raw material (in terms of stone type and form) can be pinpointed as a strong driver of outcomes: where there was higher quality (i.e., more easily knappable) raw material with accessible edges for removing flakes, Oldowan hominins could have reduced the cores more intensely and also incurred fewer errors such as step and hinge fractures (Delagnes & Roche, 2005; Gabrić et al., 2021). The same kinds of differences would therefore be observable when comparing experimental studies using different raw materials (Harlacker, 2003; Proffitt et al., 2021a).

Nonetheless, there is an identifiable pattern of lithic attributes and knapping mistakes among artifacts made by human novices, regardless of either learning condition or raw material used (Harlacker, 2003; Nonaka et al., 2010; Pargeter et al., 2020, 2021; Stout & Semaw, 2006; Stout et al., 2009; Zorrilla-Revilla et al., 2021). High frequencies of battering and step fractures, along with small, non-invasive flakes and distinctly sized platforms, are features that novice knappers from the present study share with experiments that came before, and which generally, but not necessarily, distinguish all novice knapper outcomes from the skilled work that is suspected by many authors to have been exercised in even the earliest Oldowan (see Braun et al., 2019; Stout & Semaw, 2006). The results of all these studies may represent a hypothetical learning curve that living human novices have entered via multiple avenues of learning, but which seems inaccessible to apes and which thus far appears undocumented from true Oldowan assemblages (Bandini et al., 2021c, 2022; Pargeter et al., 2021; Stout & Semaw, 2006; Stout et al., 2009; Toth et al., 2006), likely due to the aforementioned problematic site formation processes and possibilities of skill-averaging effects from the 'sharing' of materials by separate individuals and/or groups.

Being able to identify suitable platforms and strike cores from within the correct range of angles is fundamental to conchoidal fracture and, thus, flaking success (Ambrose, 2001; Cotterell & Kamminga, 1987; Delagnes & Roche, 2005; Stout & Semaw, 2006; Toth, 1985). Conchoidal fracture as the mechanical process for creating flakes (rather than just angular fragments) is central to arguments about the comparability of artifacts from living apes and humans to the Oldowan record's many artifacts and assemblages (e.g., Proffitt et al., 2016, 2018, 2022a; Stout et al., 2009; Toth et al., 2006). Therefore, the confirmation that conchoidal fracture did indeed occur during the toolmaking bouts of our participants – demonstrated by artifacts with the qualitative attributes key for flake classification, as well as EPAs well within the range from an Oldowan reference dataset (Režek et al., 2018) is support for the external validity of our results. Perhaps glass with such 'obvious' (i.e., obvious to an experienced knapper or lithicist) available knapping platforms increased frequency of flaking successes both intra- and interindividually (cf. Dogandžić et al., 2020; Luncz et al., 2022a; Proffitt et al., 2021a), but this does not mean that no flaking would have occurred among naïve individuals had they been provided other materials for potential toolmaking (again, this is likely a strong influence of observed artifact variation at different Oldowan sites; Delagnes & Roche, 2005). This assertion is verified, at least in part, by the fracture of river cobbles by several participants in our study (n = 11) and the fracture of the granite block as well (n = 5).

As previously mentioned, across the participants there was re-innovation of all four categories of early knapping techniques, which have been associated with Oldowan and pre-Oldowan toolmaking (de la Torre, 2019; Harmand et al., 2015; Putt, 2015; Toth & Schick, 2011, 2018; Wynn & McGrew, 1989). Freehand technique is the one most often typically implicated in controlled flaking behavior like that of the Oldowan and also in later periods (de la Torre, 2019; Putt, 2015; comment by Shipton on Tennie et al., 2017; Shipton, 2020). There were numerous instances of re-innovation of freehand technique (n = 20). Despite this, the relative frequency of freehand technique in toolmaking events might be interpreted as low compared to what is usually interpreted to be the case for the Oldowan (e.g., de la Torre, 2019). However, the frequency of freehand knapping in the Oldowan might be overinterpreted base on observations from improper or limited experimental assemblages. Differentiation of knapping techniques can be problematic, with considerable overlap in quantifiable traits and qualitative attributes of products from bipolar and freehand knapping (cf. Byrne et al., 2016; Duke & Pargeter, 2015; Hiscock, 2015; Pargeter & Eren, 2017). This overlap in outcomes of separate techniques becomes even more problematic if single artifacts were modified with more than one technique or gesture during their lifespans (e.g., B. Morgan et al., 2015; Hiscock, 2015; Snyder et al., 2022). An assemblage where, for example, freehand technique was used in modifying stone at early stages of reduction and bipolar knapping was used in later stages might give the appearance of involving only freehand knapping, therewith leading to an underestimation of non-freehand knapping. In the opposite direction, more 'crude' techniques like bipolar and passive hammer technique might be favored by novices with limited motor control (e.g., Duke & Pargeter, 2015; Ferguson, 2003; Geribàs et al., 2010; B. Morgan et al., 2015; Pargeter et al., 2021; Sternke & Sørensen, 2007) and then be replaced by freehand technique as their knapping skills increase (a sort of natural fixation on freehand technique as their sensorimotor coordination improves), meaning assemblages produced exclusively by novices would show less evidence of freehand knapping than could be expected from assemblages coming from populations with humans or hominins of varied skill levels. Due to these arguments (as well as the fact that behavior frequencies on the population level can be guided by equifinal cultural processes; Acerbi et al., 2022), the most important observation is that the techniques are present at all and *not* the exact frequency in which they appear.

The technology produced by the technique-naïve knappers possessed more-or-less an Oldowan-like character. The representation of the types (Leakey, 1971) suitably fits a model in which the initial blank form is highly deterministic of the potential core outcomes (Toth, 1985). Overall, all expected outcomes (defined as core types) from hemispherical, 'split-cobble' blanks were produced during this study (Toth, 1985). The reduction sequences, as interpreted from the patterns of removals from the surface of the cores (de la Torre, 2011), also were typical of Oldowan knapping 'strategies.' Consequently, our results prove further evidence of the notion that Oldowan core types and reduction sequences are most likely the produce of least-effort flaking for the efficient and/or expedient production and use of cutting edges (e.g., Moore & Perston, 2016; Toth, 1985). The most frequent core type was multifacial cores, which seemed to be most strongly linked to the reduction intensity, with multifacial cores possessing significantly less original blank mass than other core types.

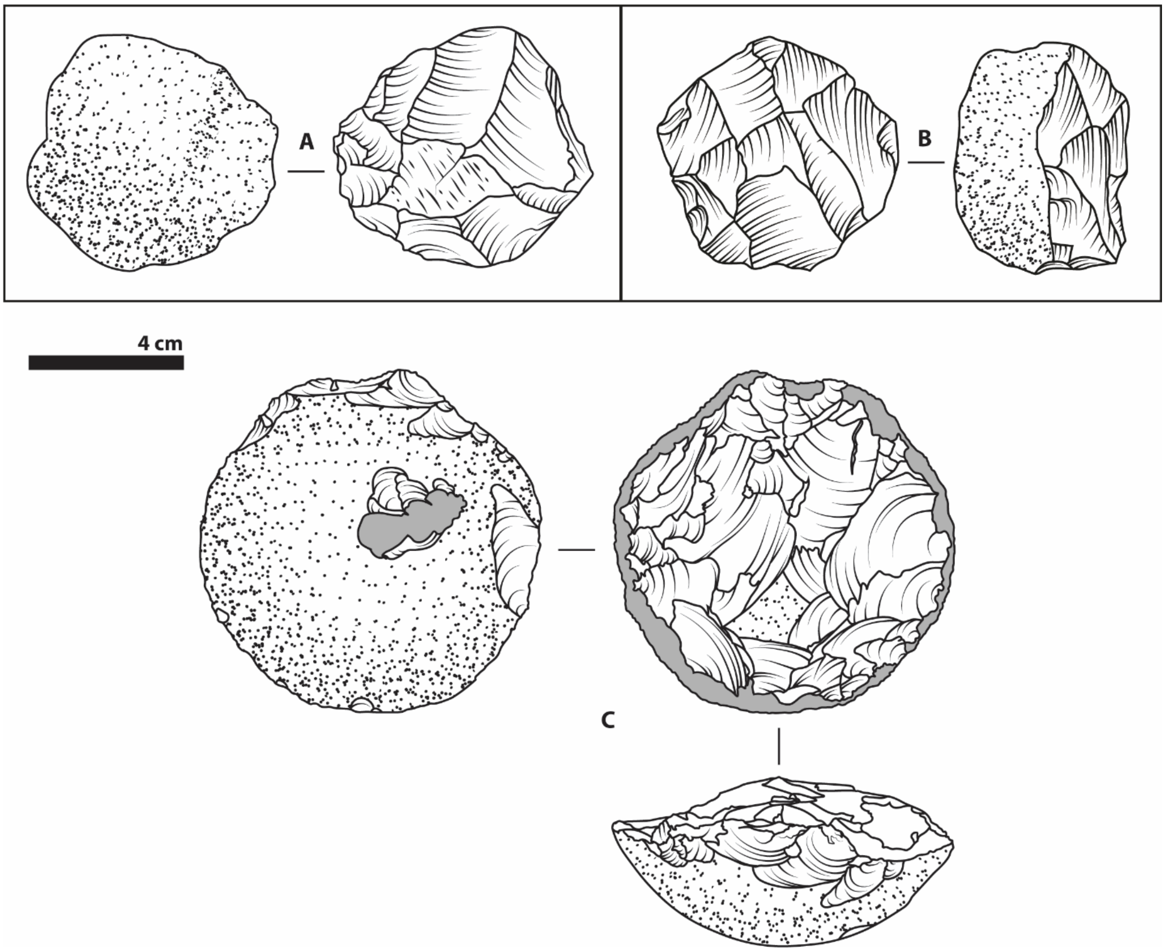


Figure 45 Radial, or centripetal, reduction. The second most frequent core morphology from this study, after multifacial cores, were radial, or centripetal, cores. Here are two archaeological examples with radial knapping sequences: a “partial discoid” from Gona, Ethiopia dated to approximately 2.6-2.5 Ma (Semaw, 2006) and “unifacial bipolar flaking” on obsidian (volcanic glass) from Melka Kunture, Ethiopia, where exploitation of obsidian during the Oldowan started around 1.7 Ma (Piperno et al., 2009). These archaeological examples are shown in comparison with a selected similar core, which was produced by a totally naïve participant in this study (P14).

Given the suitability of knapping platforms on the hemispheres, it is unsurprising that there was such high representation of centripetal cores (n = 8), which resembled unifacial or partial discoids from the earlier Oldowan (Figure 45; Semaw, 2006; Toth, 1985). There was a single bifacial centripetal core, but this did not clearly display the hierarchical reduction patterns of such cores from the later Oldowan and early Acheulean (e.g., de la Torre, 2011; de la Torre et al., 2003; Leader et al., 2018). It is unclear whether – though hypothetically possible that – such bifacial hierarchical centripetal cores might have developed with additional removals from the unifacial cores or with additional testing time for the participants to hone their skills.

Although the hemispherical cores are generally unreliable for biface production (being too small and not well-shaped), there was one curious case of a potential biface that appeared during testing of P17. The size, superficial morphology, and reduction pattern of the object (which was subsequently further reduced until becoming a multifacial core) fit the definition of a protobiface (Leakey, 1971; Shea, 2010), a fairly typical phenomenon in later (Developed) Oldowan and early Acheulean assemblages (Leakey, 1971; Leader et al., 2018). Similar to the case of Oldowan core types, this evidence supports the notion that early Acheulean bifaces might not be products of mental templates, but rather more unintentional flaking sequences (compare with the appearance of bifaces in spandrels experiments; Moore & Perston, 2016). This leads to the possibility that larger Acheulean handaxe types, like handaxes and cleavers (Ambrose, 2001; Shea, 2010), might appear in baseline knapping conditions if initial blank morphologies more typical of the Acheulean were to be used (e.g., elongated pebbles and large flake blanks).

The artifact outcomes also seem to be mostly if not completely unintentional when regarding the responses to the naivety questionnaires. Participants who made sketches of the stone tools and artifacts to which they were previously exposed always drew objects that were not of an Oldowan-like character (some perhaps could be interpreted as hand-axes, but most drawings resemble technologies from much later periods, e.g., arrows, hafted axes, and millstones).

Finally, principal component analysis of important technological variables (after Braun et al., 2019) revealed that – despite differences, as noted above – the technology of human novices in this study were generally much more similar to Oldowan and Acheulean assemblages than to the pre-Oldowan assemblage from Lomekwi 3 (Harmand et al., 2015) and stone artifact assemblages from wild capuchins (Proffitt et al., 2016). The strongest differences between the experimental and archaeological samples were in terms of the flaking scars on cores and the relative size of flakes, which would indicate relative early reduction stage of many cores as well as the relative inexperience of the human novices. Again, significant differences between flakes produced by participants in this study and flakes from Oldowan assemblages mirror previous research on human novice toolmaking, wherein novice products were relatively small and thin (Harlacker, 2003; Nonaka et al., 2010; Pargeter et al., 2020, 2021; Stout & Semaw, 2006; Zorrilla-Revilla et al., 2021). This indicates that a large portion of the observable differences between technology of human participants in this study and the known Oldowan record is more related to early learning indicators, which are not readily apparent at true archaeological sites (cf. Stout & Semaw, 2006; Stout et al., 2009).

Theoretical implications for early hominin cultures

Irrespective of the nuances outlined above, the results of this study are *strong* proof-of-principle that the central components of the Oldowan technology, early knapping techniques, the resulting artifacts, and cutting tool use, do not require any cultural transmission of know-how in order to be expressed.

While re-innovation by a single naïve individual would have been sufficient to serve as this proof-of-principle, the study on naïve humans provided significantly more cases of knapping technique re-innovation[4]. Among the technique-naïve participants (n = 25), there were a total of twenty cases of re-innovation of passive hammer technique, twenty-two cases of re-innovation of bipolar technique, twenty cases for freehand technique (especially pertinent given the suppositions by Shipton, 2020; see also comment by Shipton on Tennie et al., 2017 and), and ten cases for projectile technique (one verifiable whole flake was produced from bouts of projectile toolmaking). Because participants were all tested individually, these numerous occurrences of knapping technique re-innovation serve as independent, in-study replications and, thus, well-substantiated proof that all four categories of early knapping techniques can be individually re-innovated. The phylogenetic proximity of living modern humans to extinct hominins strengthens the validity of the proof-of-principle, while know-how re-innovation in other extant primate species (capuchin monkeys and orangutans: Motes-Rodrigo et al., 2022; Westergaard & Suomi, 1994a, 1995a) shows that indirect cultural transmission of know-how is not needed to express early knapping techniques and further reinforces the sincere possibility that re-innovation of stone tool-related know-how could have occurred in further species: pre-modern hominins.

Previously, experimental and cognitive archaeologists have conducted studies wherein modern human novices were given varied opportunities to learn via types of cultural transmission of know-how and, based on those studies' results, concluded a need for some type of copying social learning for transmitting the know-how relevant to stone toolmaking (Cataldo et al., 2018; Lombao et al., 2017; T. Morgan et al., 2015; Pargeter et al., 2021). Here, we have demonstrated that all of these mechanisms for cultural transmission of know-

[4] In this context, 're-innovation' refers to the very first use of any specific early knapping technique by a single naïve individual. A single individual can therefore be responsible for up to a maximum of four re-innovations, one for each of the four early knapping techniques studied here.

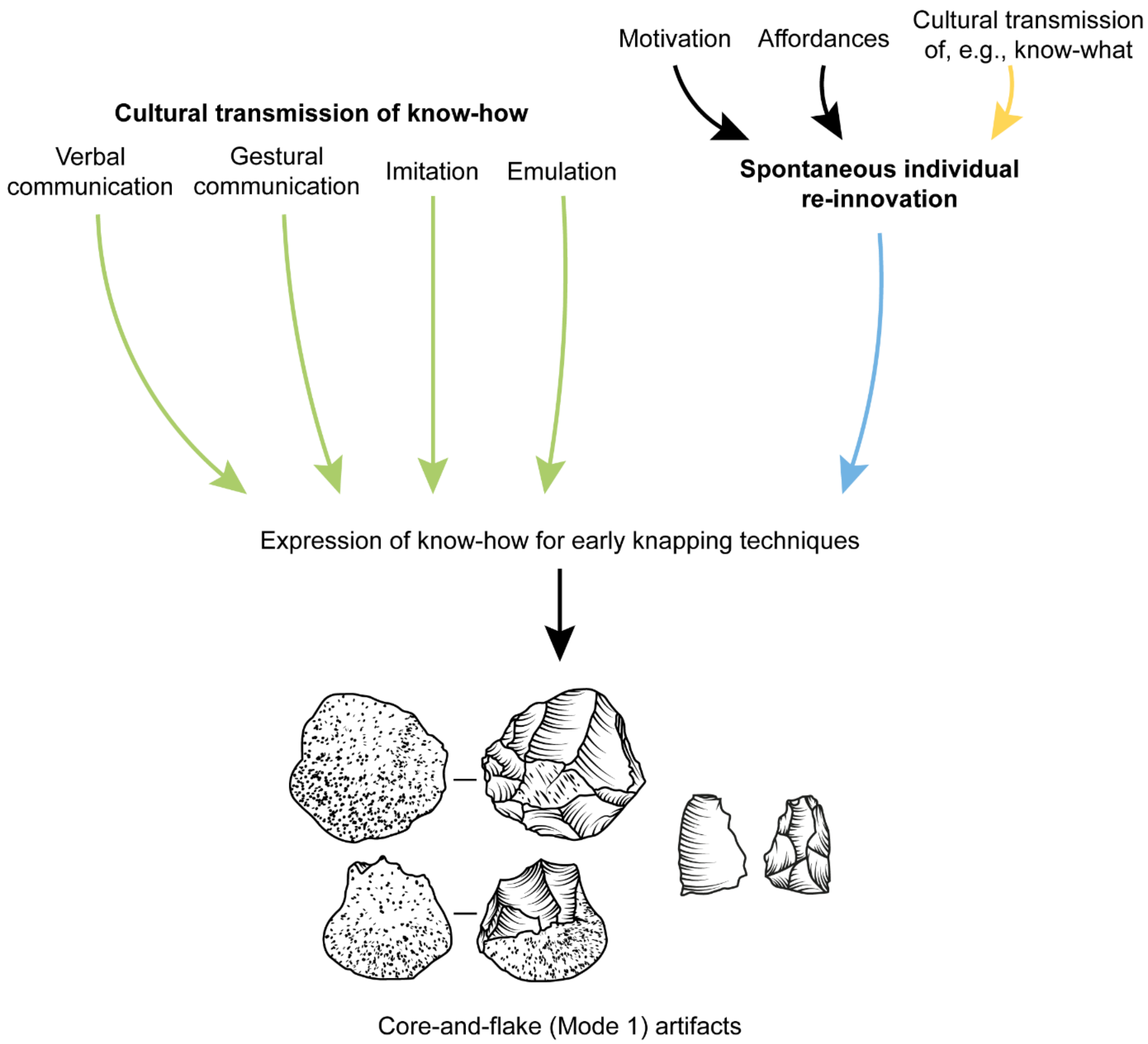


Figure 46 Equifinality of learning conditions. In novice knapping studies, many mechanisms of cultural transmission of know-how have been shown as valid means of transmitting knapping know-how to WEIRD human participants – regardless of contentious differences in efficiency – including verbal communication, gestural communication, imitation, and emulation (Cataldo et al., 2018; Lombao et al., 2017; T. Morgan et al., 2015; Pargeter et al., 2021; Putt et al., 2017). However, spontaneous individual re-innovation in island tests has also proven – in living modern humans – to be a suitable context for the development of knapping know-how (Snyder et al., 2022). As such, there is a clear case of equifinality of processes. Oldowan assemblages present us only with outcomes (here, represented by cores and a flake from Gona; Semaw, 2006) and not processes. Parsimony would favor the option requiring the fewest number of assumed evolutionary steps: thus, the re-innovation pathway and not any of the cultural transmission of know-how pathways.

how – even what some have considered to be the minimum fidelity mechanism, emulation (see T. Morgan et al., 2015; comment by Nielsen & Whiten on Tennie et al., 2017; but compare also with criticisms of experimental emulation conditions by Pargeter et al., 2021) are unnecessary for the development of early knapping techniques. Even totally naïve human participants are capable of individually learning how to knap and therewith producing Oldowan-valid artifact types. By presuming a minimum contribution of some copying social learning, these novice knapping studies have ignored the possibility of a non-copying pathway to know-how development (Tennie et al., 2016). In the end, humans are capable of requisite knapping know-how and the artifactual consequences in all thus-far tested social learning contexts, proving all of the mechanisms and processes to be equifinal; the comparative efficiency and speed of learning by living modern humans in these learning contexts are not particularly informative, because the archaeological record only presents us with process outcomes, and these may have been produced over decades or millennia as opposed to minutes or hours, as in the case of experimentally-generated artifacts and assemblages (Figure 46; see Perreault, 2019). Therefore, the best explanation for the presence of Oldowan and other Mode 1 artifacts and assemblages would be the most parsimonious of considered options: individual re-innovation, supported by cultural transmission of information types other than know-how (i.e., minimal culture: Bandini & Tennie, 2017; Bandini et al., 2020; Buskell & Tennie, in press; Neadle et al., 2017; Tennie et al., 2016, 2017, 2020).

Acknowledgments

We particularly and greatly appreciate the valuable input that we received throughout the research process from S. McPherron. Similarly, E. Bandini has played a key supporting role from the very beginning. We would also like to thank the following colleagues for advice and support: A. Blanco Lapaz, D. Boysen, N. Conard, K. Harvati, W. Hütteroth, L. Li, N. Michiels, A. Motes Rodrigo, D. Neadle, J. Orellana Figueroa, E. Reindl, and B. Starkovich. Special thanks to J. Keppeler for performing the interobserver reliability coding, to A. Falcucci for performing the interobserver reliability flake identification, and to N. M. Morales Garcia for work on Figures 19 and 20.

VII. General discussion: the minimal culture model and further considerations

Explanation of candidate contribution to collaborative work

This chapter is based on collaborative work. The text is partially adapted from the main text of the following publication:

Snyder, W.D., Reeves, J.S., & Tennie, C. (2022). Early knapping techniques do not necessitate cultural transmission. *Science Advances, 8*(27), eabo2894. doi:10.1126/sciadv.abo2894

Conceptualization: C.T., W.D.S., J.S.R. Investigation: W.D.S. Visualization: W.D.S. Formal analysis: W.D.S., C.T. Writing (original draft): W.D.S. Writing (review and editing): W.D.S., C.T.

Some of the remaining content is from as-yet unpublished, in-preparation work:

Snyder, W.D. & Tennie, C. (in press). What kind of culture did early hominin toolmakers have?. *Mitteilungen der Berliner Gesellschaft für Anthropologie, Ethnologie und Urgeschichte*.

Snyder, W.D., Reeves, J.S., & Tennie, C. (in prep). Untitled naïve human toolmaking and tool use paper.

Evolutionary processes, not just biological but also cultural, played an undeniable role in the development of the hominin lineage. Research in the 20th and 21st centuries has revealed the capacities of nonhuman animals (e.g., van Schaik et al., 2003; Whiten et al., 1999) – especially our closest primate relatives – for culture, but those cultural capacities, traditions, and technological outcomes (which we call minimal culture; Neadle et al., 2017; Tennie et al., 2020) appear distinct from the flagship component of human cultural life, the cumulative culture of know-how (Boyd, 2018; Boyd & Richerson, 1996; Henrich, 2016; Stout & Hecht, 2017; Tennie et al., 2020; Tomasello, 1999). From this perspective, we suspect that the origins of modern human-like cumulative culture of know-how should have taken place at some point during the hominin lineage (e.g., Boesch & Tomasello, 1998; Mithen, 1996; Montrey & Shultz, 2020; Tennie et al., 2016, 2017; Toth & Schick, 2018; Tramacere & Moore, 2018; Snyder et al., 2022; Stout & Hecht, 2017; Stout et al., 2019; van Schaik et al., 2019; Wynn & McGrew, 1989). Here, we outline the evidence for cumulative culture of know-how during the Oldowan, ultimately opting for a minimal culture model (Snyder et al., 2022) for the technological industry. We also explore further implications regarding technological and cognitive evolution in the Lower Paleolithic/Early Stone Age.

The stasis conundrum

Stasis is a central theme in the discussion about cumulative culture of know-how in the early stone tool record. Most researchers have tended towards the interpretation that Early Stone Age/Lower Paleolithic was a time of sheer technological invariability, whereby neither the know-how involved nor the morphology of the artifacts produced did not change much (e.g., Cueva-Temprana et al., 2022; Foley & Lahr, 2003; Gallotti, 2018; Horta et al., 2022; Isaac, 1972, 1984; Jelinek, 1977; Semaw et al., 2003; Tennie et al., 2016, 2017; van Schaik et al., 2019). Some others have gone against this trend and proposed that there is actually regional and temporal variability in the Oldowan (Barsky, 2009; Stout et al., 2019; Toth & Schick, 2011, 2018). If indeed there was a technological stasis in the early stone tool record – which most likely appears to be the case (see below) – then how did it come about that behaviors and artifact forms were 'maintained' across such long stretches of geologic time (e.g., Foley & Lahr, 2003; Mithen, 1996; Tennie et al., 2020)?

It has been suggested that raw material variables can restrict the 'wiggle room' for innovation that would have gone beyond the technological know-how identified with the Oldowan (e.g., Hovers, 2012; discussed in Tennie et al., 2020). And yet, other possibilities do exist, as is clearly demonstrated by the subsequent development of Acheulean knapping sequences along with hundreds of thousands of years of innovations in knapping sequences and techniques and radiation of artifact morphology in the Middle and Upper Paleolithic (Ambrose, 2001; van Schaik et al., 2019). In fact, with the raw materials at their disposal – in terms of blank forms and raw material types – Oldowan hominins, if they were truly so copying-savvy, should have been able to cross the 'Rubicon' to achieve more complex, hierarchical knapping sequences (e.g., as in the Levallois industry, prototypical versions of which already seemingly arising in the Developed Oldowan between 1.6 and 1.3 Ma; e.g., de la Torre, 2011; de la Torre et al., 2003). Raw material constraints alone were insufficient to explain long-term periods without major technological change, which therefore places a primacy on more biological, cognitive, or social factors for explaining these observations.

'High-fidelity' copying, such as imitation, has been implicated to explain the lengthy technological stasis in the Oldowan (discussed in Tramacere & Moore, 2018, p. 324; cf. T. Morgan et al., 2015). This implication, however, directly contradicts modelling-based predictions in cultural evolution research that intrinsically connect copying social learning with 'inevitable' copying error as a minimal expected source of know-how ratcheting and diversification (Eerkens & Lipo, 2005; Schillinger et al., 2014). But there is no true incremental change in the Oldowan. Instead, hominins seem to have been doing the same things throughout the Oldowan and also in Mode 1 (e.g., Braun et al., 2019); the variation that exists does not follow a continuous and uniform temporal pattern of technological improvement or progression (e.g., Cueva-Temprana et al., 2022). Although this remains fairly understudied, shifts that might exist within the Oldowan, e.g., as with the Developed Oldowan, may still relate more to changes in hominin biology (Pradhan et al., 2012; Stout et al., 2008; Tennie et al., 2016, 2017; cf. Stout et al., 2019; see also below). The opposite has also been posted: some accounts propose that the limitations of imitation and emulation in spreading innovations might be related to the snail's pace of technological change (T. Morgan et al., 2015), but here again these social learning mechanisms are falsely identified as low fidelity when there is empirical proof that even emulation can transmit know-how and thus lead to cumulative cultural evolution of know-how (Caldwell & Millen, 2009; Reindl et al., 2017).

Conformity bias (otherwise, conformist bias) has further been implicated as a mechanism to counterbalance the effects of copying error and other processes that generate increasing diversification of know-how and artifact outcomes (Lycett & Gowlett, 2008; T. Morgan et al., 2015; see Tennie et al., 2017). But the "faithful replication of systems" (Foley & Lahr, 2003, p. 119) shown by the early archaeological record and supposedly achieved through the suggested combination of copying social learning and conformity bias is well beyond the fidelity of copying (and thus, cognitive capacities for imitation) that is observed in living modern humans, while copying of novel know-how under conformity is beyond the capacities of apes (see Neadle et al., 2021). Furthermore, there is strong evidence to conclude that conformity bias should rather lead to an increase in intergroup or interpopulation know-how differences rather than such widespread unity within the Oldowan industry across space and time (reviewed in Tennie et al., 2020; e.g., Boyd & Richerson, 1985; Henrich & Boyd, 1998; Richerson & Boyd, 2005).

It has also been proposed that individual learning as the driving mechanism for know-how acquisition would "constrain variation somewhat, but costs associated with knapping would encourage the adoption of social learning mechanisms that would countermand the inevitable effects of copying error with reduced risk to tool manufacturers" (Lycett et al., 2015, p. 155). This interpretation relies – again – heavily on circular reasoning, as it already implies that there must essentially be cultural transmission of know-how at play from the start when it may rather only be a byproduct of biocultural co-evolution in living modern humans, while also assuming that learning costs (and benefits) and learning efficiency would be the same in extinct hominins as in extant humans (e.g., Hiscock, 2015; Lycett et al., 2015; T. Morgan et al., 2015; Pargeter et al., 2019, 2020; Sterelny, 2021; Sterelny & Hiscock, in press; Stout, 2002; Toth & Schick, 2018). Here, putative costs of learning are exaggerated, too often relying on a model where toolmaking is a separate entity from tool use where you sit and practice until you are 'good' at toolmaking (e.g., Pargeter et al., 2019, 2020; Stout et al., 2009), when actually 'bad' toolmakers can still occasionally make good tools and 'bad' tools can also be useful for something (cf. Hayden, 2015). Even assuming that individual learning is as costly as is being predicted, it does not mean it is not sufficient – along with non-copying social learning – for the development of know-how for 'simple' flaking technologies. Indeed, our experimental data support this.

Simply put, processes of individual re-innovation and non-copying social learning working in tandem is a more parsimonious explanation for long-term stases in the Lower Paleolithic (see Tennie et al., 2017). Variability occurs within this stasis (e.g., Jelinek, 1977) because of repeated re-innovations of the technology in disconnected populations, where "knowledge is repeatedly lost and rediscovered" (Montrey & Shultz, 2020, p. 6; Braun et al., 2019; Cueva-

Temprana et al., 2022; Tennie et al., 2016, 2017). Regional differences between sites are better characterized as samplings of behaviors and artifact types that are part of the larger repertoire (potentially the ZLSs of several different hominin species) of toolmaking and tool use behaviors that would have been re-innovatable by populations of early hominins. The patterns of differences (e.g., de la Torre, 2019) arise due to differential emergence of select behaviors across populations experiencing mechanisms of serial re-innovation of know-how (here, the agent-based modelling work in Acerbi et al., 2022 applies; contra Toth & Schick, 2011, 2018).

Elimination of 'non-cultural' explanations?

The method of exclusion 'determines' the action of cultural processes by ostensibly eliminating non-cultural (in this regard, biological/genetic and ecological, as well as raw material) factors (the most classic example being Whiten et al., 1999). The method of exclusion was used originally for primate studies but has also applied in some manner to archaeological research (e.g., Stout et al., 2019). Stout et al. (2019) consider cumulative culture of know-how to be the most likely explanation for inter-site differences in core rotation versus core non-rotation during toolmaking, "due to multiple unsupported assumptions that are required" for considering the role of biological and cognitive differences between hominins as well as ecological and raw material factors in said core rotation distinctions (p. 314).

The faultiness of methods of exclusion for studying ape cultures notwithstanding (see Acerbi et al., 2022; Motes-Rodrigo & Tennie, 2021; Neadle et al., 2017), we regard this conclusion to be a false premise and, indeed, itself a non-parsimonious explanation, given the broad empirical evidence from living primates for the interaction of cultural behaviors with genetic and environmental factors from primates.

Raw material can influence – also when interacting with individual skill level – both behavior and behavioral traces that primatologists and archaeologists can identify (Proffitt et al., 2021a, 2021b, 2022b). It is especially pertinent that the purposes of tool use and toolmaking can determine not just what tools are used but how they are used (e.g., Proffitt et al., 2021, 2022b). In the non-stone tool realm, ant-dipping is one example of where tool differences might superficially appear to involve cultural preferences, but instead, ant-dipping tool lengths are related to aggressiveness of the prey ant species (Humle & Matsuzawa, 2002;

Schöning et al., 2008[5]). Similar patterns might be at play at Oldowan sites, if we are to also generalize the results of our study, where the purpose of artifacts (as a core versus as a core-tool) might have an influence on the variability of types that appear and the reduction patterns (e.g., unfacial versus bifacial).

Stout et al. (2019) suggest that there are too many assumptions to consider differences in individual learning capacities as an explanation for variation in the record. This defies the almost certainty that, just as in living primate species and subspecies (Bandini & Tennie, 2018; Bandini et al., 2022; Forss et al., 2019; Gruber & Clay, 2016; Gumert et al., 2019; Herrmann et al., 2010), hominin taxa and/or populations would have varied considerably in their cognitive and cultural capacities, even in cases where the separate hominin taxa and populations would have been contemporary with one another (e.g., Delagnes & Roche, 2005; Semaw et al., 1997; Stout & Hecht, 2017; Tennie et al., 2017). In some cases, genetic differences could have led to total non-expression of a behavioral outcome (as in what appears to have in bordering subspecies of long-tailed macaques; Bandini & Tennie, 2018; Gumert et al., 2019), while in others, this would explain general propensity towards certain behaviors, the frequency at which re-innovations might occur, and slight differences in the form of behaviors as a consequence of being channeled by tweaked cognitive mechanisms (Forss et al., 2019; Herrmann et al., 2010; Tennie et al., 2020).

Differences in tool use and toolmaking strategies might also appear – or become more exaggerated – if distinct contemporary hominin taxa were under the influence of niche partitioning. Here, niche partitioning would have led to the separate groups engaging in different technological behaviors or variants of technological behaviors appropriate for their individual niches (see Dusseldorp & Lombard, 2021, regarding a similar interpretation of the sparse material culture of *H. naledi* compared to the rather diverse material culture of the co-existing early *H. sapiens*) and occupying neighboring loci with uniquely available resources that do or do not require tool use or varying forms of tool use (or similar tool use for different purposes; Proffitt et al., 2021a).

[5] A third study by Koops et al. (2015) claimed to disprove this by studying ant species availability and tool length in neighboring chimpanzee groups, finding statistically significant differences in tool lengths between the groups despite the same availability of prey. This study only measured, however, the prey selection by one of the groups (and not in the other), thereby failing to separate between prey availability in both territories and foraging preferences that might actually differ from the ecological availability. Differences in tool length depending on prey species in the one population of chimpanzees was found to be non-significant and so the authors eliminated this as a factor in determining tool length, and yet, the *p*-value was 0.085 with a non-parametric Mann-Whitney U test, which is still a rather notable result.

Just as evolutionary processes resulted in mosaic combinations of anatomical adaptations and constraints among hominin populations (e.g., Foley, 2016; Hublin et al., 2017), they should have also resulted in similar patterns in cognitive traits (e.g., Foley & Lahr, 2003). It is not so much that knapping techniques are predicted to be unchanging 'instincts' (see Bandini & Tennie, 2018; Bandini et al., 2021a), but rather that the diversity of cognition in hominins would have channeled know-how developments in such a way that the same technological (Mode 1) principles were expressed in so slightly non-identical ways, measured in behavioral and artifact forms (Proffitt et al., 2021a, 2022b; Tennie et al., 2020). In other terms, early knapping techniques might have been within the re-innovation capacities of different – likely many – species throughout time and space, but the exact manifestation of the know-how would then be filtered through various variables related to species-specific cognition, ecology, anatomy, and combinations of these and other variables. We probably should not expect, for instance, that a late australopithecine (Semaw et al., 1997; Stout et al., 2019; Toth & Schick, 2018), living modern humans (as in this study), and *H. floresiensis* (Moore & Brumm, 2009; Moore et al., 2009) would perform identical actions and make identical artifacts, but the requisite know-how and the ability to acquire it would be a core conserved trait(s).

Finally, the promotion of cumulative culture of know-how in the Oldowan is itself not parsimonious (contra Stout et al., 2019) due to the lack of copying abilities in apes and other primates (e.g., Clay & Tennie, 2018; Neadle et al., 2021; Tennie et al., 2012), the apparent relative stasis in technology (see above), and the demonstration that the early knapping techniques requisite for Oldowan toolmaking are capable of being re-innovated by naïve individuals in not just one but a few living species, including living modern humans, the closest phylogenetic relatives of the Oldowan toolmakers (Motes-Rodrigo et al., 2022; Snyder et al., 2022; Westergaard & Suomi, 1994a, 1995a). Once re-innovated, the expression of know-how could have spread within hominin populations by some process *other than* cumulative culture of know-how (see Acerbi et al., 2022).

Other hypothetical possibilities

There is any number of possibilities that could be applied to the current evidence. We only comment on two here: a) cumulative culture of know-how was present in the Oldowan but did not exceed individual re-innovation levels in complexity or was not related to early knapping techniques, and b) artifact types or knapping techniques were under Baldwinian selection and perhaps even by the Oldowan had become secondarily genetically predetermined behavior.

Firstly, one could argue that there was still transmission of know-how. Perhaps the 'transmitted' know-how had simply not yet accumulated to the degree that it exceeded the limits of what was individually re-innovatable. While this is possible, it is questionable to call this know-how transmission – as a better term for such learning would be "triggering" (compare Buskell & Tennie, 2022; Tennie et al., 2020b). On the artifact level, it would be impossible to distinguish between serially reinnovated know-how and copied know-how that is within individual reach of re-innovation, and so from this perspective, the possibility copying occurred is impossible to eliminate completely but ultimately not a parsimonious explanation of the given evidence (see above). It is possible that other stone tool-related behaviors could be candidates for copying-dependent traits, such as those related to the "procurement, selection, transport, and curation" of raw materials (Gabric et al., 2021, p. 25). However, these behaviors would theoretically be better explained by cognitive mechanisms related to planning (executive function) and local or stimulus enhancement (see also Reeves et al., 2021), or in the same sense of 'step-wise traditions' that are known from non-human animals like pigeons (Tennie et al., 2009). Material and tool transport also occurs in apes, with hominin material transport mainly standing out – quantitatively – in the distances travelled (see Reeves et al., 2021; Toth & Schick, 2009, 2018; Toth et al., 1993; Wynn et al., 2011).

Alternatively, perhaps cumulative culture was present at the time in hominins but in some other domain (e.g., Pradhan et al., 2012). the organic tool repertoire of early toolmaking hominins is largely hypothetical, with the general assumption that they would have possessed a similar material culture to living apes (e.g., Ambrose, 2001; Bandini et al., 2022; Gabrić et al., 2021; Haslam et al., 2009; Hovers, 2012; Rolian & Carvalho, 2017; Toth & Schick, 2009). There is only limited preservation of organic tools during the Oldowan, consisting of a few possible bone excavating tools (see Backwell & d'Errico, 2008; d'Errico & Backwell, 2003). Current evidence for organic digging tools indicates individual re-innovation capabilities in chimpanzees, which might also have been true of hominins (see Motes-Rodrigo et al., 2019; see also re-innovation of bone tools by capuchins: Westergaard & Suomi, 1994b). A wide number of island test experiments have further shown that ape-typical tool use behaviors can re-innovated (Bandini & Tennie, 2017, 2019; Bandini et al., 2020; Kitahara-Frisch & Norikoshi, 1982; Westergaard & Suomi, 1993). Similarly, bamboo tools have been shown re-innovatable by capuchins (Westergaard & Suomi, 1995b), which is at least one hypothetical Lower Paleolithic technology that has often been discussed with regard to the (previous) lack of evidence for Acheulean technology in eastern Asia (Schick, 1994). Overall, the suspicion that cumulative culture of know-how might have been present in the organic tool material culture of hominins but absent in the stone tool domain is not logical, considering the evidence from living primates (see also ant-dipping example above).

Exiting the technological domain, communication and proto-language might also be candidates for cumulative culture dependent behavioral/cognitive capacities/know-how, but these are hard, nigh impossible, to trace. Cognitive archaeological research have previously tried to draw conclusions about the emergence of language as it relates to stone tools (e.g., Cataldo et al., 2018; T. Morgan et al., 2015; Stout et al., 2008), but again this relies heavily on circular reasoning about the presence of such abilities (Tennie et al., 2017). Non-human research would suggest that proto-language/early language evolved for purposes rather than the transmission of technological information or know-how (e.g., Mine et al., 2022). In this case, proto-language may not even have been an efficient (even an inefficient) medium of know-how transmission (see arguments of Snyder & Tennie, in prep). Even further, gestures and calls in apes have not been shown to have a signal for copying-dependency (see Byrne et al., 2017; Motes-Rodrigo & Tennie, 2021; Tomasello et al., 1997).

Another possibility is that the know-how underlying early knapping techniques was initially outside individual's reach, i.e., that it had been the outcome of previous cultural evolution of know-how (see Corbey, 2021; Corbey et al., 2015; Hecht et al., 2015; Pradhan et al., 2012; Stout & Hecht, 2017). In this case, it would have had to be culturally transmitted for *some period of time*, but then, gradually, this useful behavior could have eventually become subject to Baldwinian selection, resulting in individuals of later generations being eventually capable of producing knapping behavior without the necessity to copy its underlying know-how (Corbey, 2021; Pradhan et al., 2012). This has been used as an explanation of the stasis in the Acheulean (Corbey, 2021; Corbey et al., 2015) and though generally not considered for the Oldowan, may still be a possibility for that technology as well. The Baldwinian selection hypothesis also infers a certain amount of fitness benefits (cf. Stout, 2006; Stout & Hecht, 2017) in order for that selection to act on learned behavioral traits, not to mention the varied interpretations of habituality of tool use and toolmaking before, during, and after the Oldowan (Shea, 2017). New entheseal evidence (Kunz et al., 2022) might indeed support the notion of tool use that was more habitual in the pre-Oldowan than in modern apes, but this indicates only the frequency of behavioral expression in individuals without necessarily pointing to any one cultural process. It is certainly likely that exaptive and Baldwinian processes were at play during hominin evolution (cf. Tennie et al., 2016, 2017), but an account that assumes the know-how was copied at some point, presumably early after its initial innovation, means that it is non-parsomonious in the sense of inferring evolutionary steps or stages that are presently not documented by any measurable traces or observable material record.

A minimal culture model for Oldowan hominins

Island tests are powerful as a discriminatory tool for determining the necessity of transmission mechanisms for the expression of target behavior (Bandini et al., 2020, 2022; Tennie et al., 2016, 2020; Tomasello, 1999). Here, we applied the island test approach to the question of baseline abilities for the development of knapping know-how in living modern humans (Snyder et al., 2022). The application of said approach is intended only as a means of variable control in order to look at the individual level of learning and thus the atomic level of know-how expression in minimal culture and is not intended to infer by extension that every Oldowan hominin lived in some kind of island test conditions.

There is no doubt that hominins would have been social creatures, so the closest that they are likely to have come to island test conditions is in those occasions where knapping had not yet been – minimally once – innovated in independent groups or populations (Tennie et al., 2016). Nonetheless, it cannot be assumed that the know-how of early knapping techniques was culturally transmitted in those contexts where knapping present (i.e., in group or populations of hominins where the stone toolmaking behaviors were already being expressed; cf. Acerbi et al., 2022; Buskell & Tennie, in press). Just as has been generally accepted to be true (e.g., "all stone knapping is cultural behavior", Stout & Semaw, 2006, p. 308), knapping by pre-modern hominins in the Oldowan was most probably cultural behavior, but cultural behaviors driven and catalyzed by processes, excluding those which support cumulative culture of know-how (cf. Tennie et al., 2020a).

Similar to what has been proposed for non-human primates (see Buskell & Tennie, in press; Neadle et al., 2017; Tennie et al., 2020), we propose a minimal culture model for Oldowan hominins. Under the model, social learning would have been present (i.e., pre-modern hominins as with living primates and humans would have had some capacities for social learning) but would not have transmitted know-how. Instead, other information types like know-what, know-when, and know-where would have been culturally transmitted via non-copying social learning like stimulus and local enhancement (Tennie et al., 2016, 2017; see also Mithen, 1996 on the Clactonian industry). Expression of know-how would have then appeared only subsequent via a process of triggering (Buskell & Tennie, in press). Rather than direct copying being responsible, the 'spread' of behaviors would have been an illusion caused by non-copying social learning's impact on the frequencies and speed of serial re-innovations of know-how within and across populations (see Acerbi et al., 2022; Bandini & Tennie, 2017; Buskell & Tennie, in press; Montrey & Schultz, 2020; Tennie et al., 2020a). Additionally, know-how would have become stabilized at both the individual- and population-level, so long as said know-how was reasonably rewarding and not otherwise hindered by other factors (e.g., costs, dangers).

In the specific example of pre-modern hominins, information like know-what (e.g., knappable stone types) and know-where (e.g., whereabouts of carcasses) would have been fundamental to the emergence of re-innovation opportunities and occurrences. The transmission of these types of information (e.g., the association of conspecifics with localities where knappable stones are available) would then have resulted in subsequent – hence, serial – re-innovations of knapping know-how in other individuals of a group. At this stage (i.e., after the initial re-innovation(s)), individuals would have continued in their toolmaking activities, with re-application of the know-how repeatedly in the same or even in different contexts allowing for the know-how to continuously develop. This process of repetition (e.g., making a flake, cutting flesh, making a flake, cutting flesh, and so on) over years would ultimately lead to improvements in an individual's skill as well as improvements on the average skill level of the population. Here, there would have been no absolute necessity for the copying of any specific know-how, because general proximity to locations, raw materials, and extractive foraging motivation (see discussion of idealness of raw material sources and outcrops for learning; Ferguson, 2003; Shelley, 1990; Stout & Semaw, 2006), as well as opportunities to engage with all of those stimuli, would have sufficed for the expression of group-level phenomena.

Stasis in the Oldowan occurred due to a lack of cultural transmission of know-how, which prevented successive ratcheting of innovations and/or minimal copying error from producing a radiation in artifact forms (Eerkens & Lipo, 2005; McElreath, 2010; Montrey & Shultz, 2020; Tennie et al., 2020a, 2020b; van Schaik et al., 2019). The limitations of individual learning and non-copying social learning on form variability would then have had clear consequences for the evolution of technology (cf. Lycett et al., 2015; Montrey & Shultz, 2020).

It has previously been asserted that the Oldowan originated via multiple independent 'inventions' (Braun et al., 2019; de la Torre, 2019; Hovers, 2012; Shea, 2017). The minimal culture model is the most suitable framework for explaining how these multiple independent invention events could have occurred, gaining additional credence from the 're-appearance' of similar technologies in later time periods, even after more advanced technologies were already in existence (e.g., Clark, 1969; Foley & Lahr, 2003; Shea, 2013; Tennie et al., 2017).

Both the emergence of and the variability in Oldowan (see Toth & Schick, 2011, 2018) and later Mode 1 technologies is best explained by raw material differences (Delagnes & Roche, 2005; Foley & Lahr, 2003; Gabrić et al., 2021; Proffitt et al., 2021b), biological and cognitive differences between hominin groups (e.g., Bandini & Tennie, 2018; Bandini et al., 2022; Brown & Gathogo, 2002; Dusseldorp & Lombard, 2021; Foley & Lahr, 2003; Forss et al., 2019; Gruber & Clay, 2016; Gumert et al., 2019; Herrmann et al., 2010; Stout & Semaw, 2006; Stout et al., 2009), differences in tool utility/application and variation in hominin niches (cf. Hayden, 2015; Proffitt et al., 2021a), differences in site use/visitation/habitation (Braun et al., 2019; Shelley, 1990; Toth & Schick, 2018), and certainly interactions thereof. Fixation of toolmaking preferences on the individual or group level would have occurred by means other than know-how copyinhg (cf. Gabrić et al., 2021; Geribàs et al., 2010; Rein et al., 2014), while, e.g., disinterest in or lack of exploitation of specific raw material types that were nonetheless available might relate to consequences of stimulus enhancement or some other non-copying social learning (cf. Gabric et al., 2021; Mithen, 1996; Tennie et al., 2017). Stochastic factors would have also been at play, especially given the role of least-effort flaking in Oldowan and Mode 1 technologies (Isaac, 1972; Moore & Perston, 1985; Toth, 1985), though often not given much attention in discussions of cultural propensities of hominins.

When weighing the available data, it becomes clear that Oldowan artifacts and assemblages are not viable as unequivocal evidence for cultural transmission of know-how during the toolmaking activities of premodern hominins. Claims for the early appearance of cumulative culture of know-how in the hominin lineage are therefore dubious, meaning that the search for the origins of modern human-like cultural capacities and cultural phenomena would be

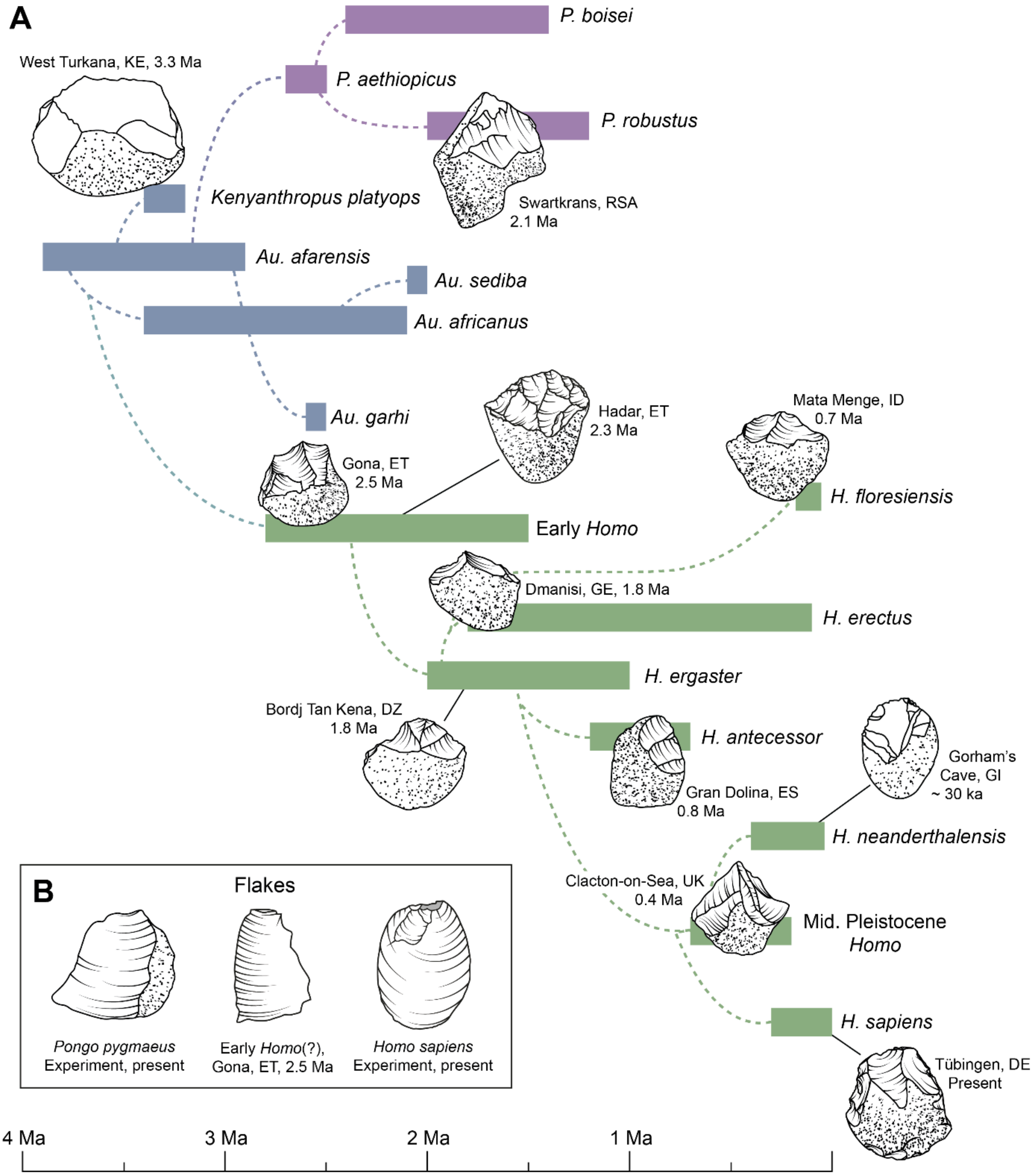


Figure 47 A timeline of hominin species showing approximate phylogenetic relationships and emphasizing the (re-)occurrence of Oldowan and Mode 1 technologies across taxa and throughout time. Illustrations based on artifacts from the identified sites (Carbonell et al., 1995; Harmand et al., 2015; Mgdeladze et al., 2011; Kuman et al., 2018; Mithen, 1996; Moore & Brumm, 2009; Motes-Rodrigo et al., 2022; Sahnouni & van der Made, 2009; Semaw, 2006; Snyder et al., 2022).

better directed at technologies appearing later than the Oldowan. Of course, there is an element of equifinality to be considered here (artifacts produced in a grey zone of cumulative culture of know-how would theoretically be undifferentiable from those made under minimal culture conditions). Of the available options, however, the minimal culture model remains the most parsimonious for explaining the archaeological record as we know it (see also comment by Tennie on Stout et al., 2019).

The evolution of cognition and technology

The impressively long record of stone artifacts is thought to be useful for tracing not just evolution of technology itself but also biological and cultural evolution of the hominins themselves (e.g., Stout, 2006; Stout & Hecht, 2017; Tennie et al., 2017; Wynn & Coolidge, 2016). In studying stone artifacts, the goal of archaeologists and paleoanthropologists is quite often to elucidate our origins and to come to a better understanding of what makes humans unique as a species and when and how that 'special' humanness came to be (e.g., Eren et al., 2020; Foley, 1987; Stout, 2002; Stout et al., 2009). Flaked stone *tools* are a seemingly unique adaptation in the hominin lineage (though see, especially, data from tool behaviors of capuchins and macaques: Proffitt et al., 2016, 2021a, 2022a; Westergaard & Suomi, 1994a, 1995b; see also orangutan data; Motes-Rodrigo et al., 2022).

Origins of stone knapping behavior

Percussive behaviors for the processing of nuts and other food items with stone or organic implements (e.g., hammers and anvils) have developed in a range of phylogenetically separated primate taxa, including chimpanzees (e.g., Bandini et al., 2022; Koops et al., 2010; Mercader et al., 2002; Neadle et al., 2020; Proffitt et al., 2022b), capuchin monkeys (e.g., Luncz et al., 2022a; Proffitt et al., 2016; Visalberghi, 1987), and macaques (Proffitt et al., 2021a, 2022a), with further examples from captive bonobos (e.g., Neufuss et al., 2016) and captive orangutans (Bandini et al., 2021b; Motes-Rodrigo et al., 2022).

In the exercise of nut-cracking by chimpanzees, accidental fracture of anvils can result in the production of sharp-edged stone, but the resultant sharp-edged stone (according to Proffitt et al., 2018, these represent only 'angular fragments', since the fracture has been deemed non-conchoidal) is not used for any subsequent cutting tasks (Mercader et al., 2002). Percussive behaviors in other primates, capuchin monkeys and long-tailed macaques, can also result in fracturing of stone (Proffitt et al., 2016, 2022a). In these cases, these fracture events occur often enough to result in full stone artifact assemblages, of which the flakes and cores fall quantifiably in the range of Oldowan stone tool variation (Proffitt et al., 2016, 2022a) and of which the overall technological character resembles more Oldowan tools than the artifacts from Lomekwi 3 (Braun et al., 2019). Again, however, neither the capuchin monkeys nor the long-tailed macaques use the flakes for any cutting actions. Additionally, the cultural context of these wild primate behaviors (i.e., whether there is any cultural transmission of know-how) is not clear from just the live observations and artifactual analysis (cf. Bandini et al., 2022; Motes-Rodrigo & Tennie, 2021). In controlled conditions, orangutans can re-innovate passive hammer technique and separately re-innovate cutting tool use (Motes-Rodrigo et al., 2022), whereas chimpanzees failed to re-innovate either toolmaking or cutting tool use (Bandini et al., 2021c). Even enculturated bonobos after years of training fail to show knapping abilities on par with Oldowan toolmakers, instead perform quite distinctly to both human and hominin toolmakers (Schick et al., 1999; Stout & Chaminade, 2007; Stout et al., 2009; Toth et al., 2006). Though capuchins are apparently capable of re-innovating all early knapping techniques and cutting tool use (Westergaard & Suomi, 1994a, 1995a), they are much more phylogenetically distant from humans and unlikely to better inform on pre-Oldowan hominin cognition than apes can (see Bandini et al., 2022).

These data would suggest that – although some basic cognitive and behavioral prerequisites are met by non-human apes – there must have been evolutionary developments in cognitive and/or noncognitive traits after the human and *Pan* lineages diverged (Pradhan et al., 2012; Stout, 2005; Stout & Hecht, 2017; contra Whiten et al., 2009b), in order to clarify the appearance and subsequent continuity in Oldowan technology a few million years later. In the scope of the minimal culture model for Oldowan hominins – and supported by evidence of re-innovation of early knapping techniques in living modern humans (Snyder et al., 2022) – the key differences in cognition may likely be unrelated to mechanisms of social learning and cultural transmission.

For instance, fine-tuning of motor control and coordination (e.g., Hecht et al., 2015; Putt, 2015; Stout & Hecht, 2017; Wynn & Coolidge, 2016) with corresponding anatomical changes to the hand (Ambrose, 2001; Karakostis et al., 2021; Kivell, 2015; Williams et al., 2010; Williams-Hatala et al., 2018, 2020) form one potential prerequisite for improved and more consistent toolmaking and toolmaking outcomes (though primates already have *some* capacities for relevant gestures and techniques: Wright, 1972; Toth et al., 1993, 2006, Toth & Schick, 2018; Westergaard & Suomi, 1994; Mercader et al., 2002; Proffitt et al., 2016, 2018, 2022a; Motes-Rodrigo et al., 2022). Similarly, anatomical changes like bipedalism (Foley & Elton, 1998; Foley & Gamble, 2009) represented a major removal of constraints to allow changes in the mind-body interface that would allow improved hand-eye coordination as well as enabling longer-distance raw material transport among other activities (Toth & Schick, 2018; Wynn et al., 2011). Just as well, anatomical changes introduced constraints, such as a reduction in canine size and upper body strength, that would drive hominins towards more habitual tool use and the invention of 'artificial teeth' (see Bandini et al., 2021c, 2022; Shea, 2017; Stout & Semaw, 2006). Here, pre-existing know-how also has an effect on the likelihood of innovations (e.g., food can be processed with the hands or teeth, or a broken cobble is a poor hammer and therefore useless; see Bandini et al., 2022; Tennie et al., 2020).

Otherwise, selective pressure on functions like working memory, spatial cognition, and causal inference (de Beaune, 2004; Placi, 2022; Read et al., 2021; Snyder et al., 2022) could also have been (partly) responsible for differences in toolmaking abilities of apes and hominins, especially with regard to the lack of subsequent tool use following sharp edge creation by primates (Davidson & McGrew, 2005; Mercader et al., 2002 Proffitt et al., 2016, 2018, 2022a). Early evidence for this latter ability not found in apes is (potentially) the discovery of cut marks at the site of Dikika dating to around 3.3 Ma, though if valid, this might have been use of naturally-occurring cutting edge rather than of a cutting tool made by a hominin (McPherron et al., 2010). Other potential explanations for the origins of toolmaking and the relatively constrained toolmaking abilities of apes relate to major changes in food preferences (e.g., marrow processing; cf. Guerbuez & Lycett, 2021), ecological changes (Pradhan et al., 2012), and changes in life histories and lifeways (Sterelny, 2021).

In terms of the multiple origins of Oldowan toolmaking, it is rather unlikely that a single cumulative cultural evolution pathway from, say, chimpanzee-like nut cracking to freehand stone knapping can be adopted. Nut cracking and other pounding behaviors are not a totally inappropriate model for a scenario where hominins could and would have discovered raw material affordances for knapping (see plentiful discussion in Bandini et al., 2022; Bril et al., 2015; Carvalho & McGrew, 2012; Carvalho et al., 2008, 2013; Haidle, 2010; Putt, 2015; Rolian & Carvalho, 2017; Sayers & Lovejoy, 2008; Williams-Hatala et al., 2018), but it is possible just as well that other behaviors (performed by the same or different species at different places and times) were also involved in scenarios of affordance discovery (such as free-hand hitting in gorillas or projectile behaviors in baboons and chimpanzees; see Beck, 1974; Hamilton et al., 1975; Masi et al., 2022; Putt, 2015; Toth et al., 1993).

Cognitive implications of the Oldowan

Knapping would still presumably be a learned behavior, but the cultural cognition of Oldowan hominins probably was more similar to that of apes than anything indicative of or necessary for cumulative culture of know-how (Tennie et al., 2016, 2017; also compare with, e.g., Wynn & McGrew, 1989; Wynn et al., 2011). According to the minimal culture model of Oldowan toolmaking – in the absence of copying social learning and cultural transmission of know-how – other cognitive mechanisms (e.g., working memory, spatial and causal reasoning, motor coordination: e.g., Haidle, 2010; Pradhan et al., 2012; Putt et al., 2017, 2019; Read et al., 2022; Stout & Chaminade, 2007; Stout et al., 2011, 2014, 2015; Tennie et al., 2017) would have played a proportionally larger role in the manifestation of core and flake technologies, as are known from the archaeological record.

Oldowan knapping has been interpreted as involving intentional, goal-oriented actions (Bril et al., 2010; Geribas et al., 2010; Stout & Hecht, 2017; Wynn, 2002). While this may still be the case for most behavioral traces that were left by toolmaking hominins (cf. Delagnes & Roche, 2005; Shea, 2017), there is the distinct possibility that a proportion (even a meager proportion at that) of events in primarily intentional knapping product accumulations might have been ‘accidents’ and that even some sites might be only extensive conglomerations of unintentional knapping events (akin to what occurs in capuchins and long-tailed macaques: Proffitt et al., 2016, 2022a; keeping also in mind the analyses in Braun et al, 2019). Affordance discovery episodes, as apparently occurred in naïve human toolmakers (Snyder et al., 2022), lend credence to the idea that this would have contributed partially to early lithic assemblages. Intentionality on the scale of core morphologies and reduction sequences can be – for the most part – excluded from discussions on Oldowan technology. Earlier research has already shown quite conclusively that Oldowan artifact types are most probably the byproducts of least-effort flaking (Isaac, 1972; Toth, 1985), with valid Oldowan (and even Acheulean) types appearing from purely stochastic knapping in spandrels experiments (Moore & Perston, 2016). Although there was an observable relationship between core type variability and tool use preferences in our study, the pattern was not statistically significant, and it is difficult to generalize the pattern to Oldowan material.

The basic requirements of conchoidal flaking are quickly learned, but reliable production of large flakes and controlled core reduction require more substantial investments in perceptual-motor skill acquisition. (Apel, 2001; Stout & Khreisheh, 2015; Stout et al., 2009; Pargeter et al., 2021). Central to this is the ability to discern about space and causation (e.g., De Beaune, 2004; Nonaka et al., 2010), as well as to manage the action parameters and physics involved in the flaking procedure (Bril et al., 2010; Braun et al., 2019; Delagnes & Roche, 2005; Stout & Semaw, 2006; Stout et al., 2010). Our results tend to agree with this notion (i.e., that the basic flaking abilities are acquired quickly and easily), while also disproving assertions by other authors that totally naïve individuals would not be able to develop suitable gestures for achieving flake removal (Lombao et al., 2017; Shipton, 2020). During learning, where sensorimotor deficits occurred, hominins could have innovated alternative strategies and techniques – even temporarily, as a part of the learning procedure – in order to achieve conchoidal fracture and flake production via easier or more expedient pathways (e.g., Ferguson, 2003; Rein et al., 2014; Sternke & Sørensen, 2007).

Even the earliest Oldowan knappers are interpreted to be skilled masters of intentional flaking (Braun et al., 2019; Delagnes & Roche, 2005; Stout & Semaw, 2006; Stout et al., 2019), with researchers identifying refined versus unrefined skills at different sites, as well as supposed group-typical knapping gestures and reduction sequences (Delagnes & Roche, 2005; Stout & Semaw, 2006; Stout et al., 2009, 2019). The refined skills at some early Oldowan sites (Stout & Semaw, 2006) has been the basis of argumentation favouring the necessity of copying social learning. But this may unnecessarily underestimate the extent of skill acquisition that is and can be achieved with only individual learning processes.

By any account, individual learning is absolute prerequisite for the acquisition of knapping skills. Whereas conceptual knowledge seems rather unimportant to a simple flaking paradigm (Harlacker, 2003; Snyder et al., 2022), the ability and opportunity to engage with the materials oneself is fundamental to developing an understanding of the affordances and mechanics of raw materials and then refining motor skill by 'practice' (Harlacker, 2003; Lombao et al., 2017, Stout & Hecht, 2017; Stout & Semaw, 2006, Stout & Chaminade, 2007; Stout et al., 2011). In models based on inferences of cultural transmission of know-how, this is suggested to involve intentional and extended practice (e.g., Stout & Hecht, 2017), but in the minimal culture model, opportunities for individual learning are part of a fluid sequence of toolmaking and tool use. In the latter case, the status of toolmaking being an outcome-oriented behavior is emphasized: the exact gestures and techniques do not matter (toolmaking is *not* like dancing where the steps are to be pristinely copied), nor does the precise quality of every single tool matter. Sometimes even tools that modern archaeologists interpret as 'poor quality' can have utility (Hayden, 2015) and when they do not have utility, provisioning of food resources by group members could have outweighed the costs of failed toolmaking bouts (contra, e.g., Pargeter et al., 2019, 2020).

Finally, there has been some discussion of other cognitive properties in relation to the Oldowan. This includes executive functions like foresight and planning (related to raw material transport and not so much to reduction sequences; Delagnes & Roche, 2005; Haidle, 2010), as well as working memory processes, which are implicated much more often in the late Acheulean rather than for least-effort flaking behaviors in the Oldowan (e.g., Moore & Perston, 2016; Wynn & Coolidge, 2016).

Major transitions and (dis-)continuity

In large part, evolution of technology during the Lower Paleolithic is characterized by punctuated equilibrium (more akin to expectations of biological evolution than that of true cumulative culture). One of the ‘punctuations’ on that timeline of technological evolution is the transition from the Oldowan to the early Acheulean. This transition is generally considered to have been significant (cf. de la Torre, 2016), but there are many different interpretations. For one, it should first be noted that Oldowan and Acheulean toolmaking extensively overlapped (meaning the transition was not a sudden event and that often the same species were producing both Mode 1 and Mode 2 technologies; see Duke et al., 2021; Lepre et al., 2011; Semaw et al., 2009, 2020).

Oldowan toolmaking is typically regarded as relatively simple, with the main technological know-how being the production of flakes for use as cutting tools (Isaac, 1972; Toth, 1985). In the early Acheulean, it is believed that new ‘innovations’ in know-how are incorporated into the cognitive and behavioral toolkit of premodern hominins, including the production of large flake blanks for invasive and bifacial retouching and the intentional shaping of symmetrical core-tools based on mental templates (Ambrose, 2001; Foley & Lahr, 2003; Stout et al., 2015; comment by Wynn on Tennie et al., 2017). The former aspect is universally accepted, while the latter – intentional shaping based on mental templates – is often debated, especially for early Acheulean tool forms (cf. Iovita & McPherron, 2011; Moore & Perston, 2016; Tennie et al., 2016, 2017). It is unclear whether shaping was present in the early Acheulean or if it appeared earlier (in the late or Developed Oldowan; Duke et al., 2021) or later (i.e., in the late Acheulean).

Just as with the Oldowan, the Acheulean is associated with an extensive period of stasis wherein "samples of equivalent ages vary minimally within and between major regions" (Shea, 2017, p. 210), resulting in a paradox that researchers interpret with different models: genetic predetermination of handaxe shape as a result of Baldwinian selection (Corbey, 2020; Corbey et al., 2016), latent solutions (Tennie et al., 2016, 2017; see also earlier interpretation by Davidson & McGrew, 2005), and cultural transmission of know-how and cultural determination of Acheulean forms (Foley & Lahr, 2003; Lycett & Gowlett, 2008; Schillinger et al., 2015; Shipton, 2010, 2020; Shipton & Nielsen, 2015).

Many factors have been put forth as being involved in the transition from least-effort flaking to more explicit, hierarchical knapping sequences typical of the Acheulean (Stout et al., 2015, 2021). Archaeologists have implicated cognitive changes in the Oldowan-early Acheulean transition (e.g., Foley & Gamble, 2009; Gabora & Steel, 2020; Mithen, 1996; Wynn, 1988), with some identifying the Acheulean as the first technology that is truly distinct to what apes can do (Foley & Lahr, 2003; Whiten et al., 2003). The Acheulean is claimed to have involved improvements in (copying) social learning on what was – previously assumed to have been – present in the Oldowan (e.g., T. Morgan et al., 2015; Pradhan et al., 2012; Shipton, 2010). The transition into the early Acheulean coincided with an increase in brain size (e.g., Stout & Chaminade, 2007), as well as a shift towards modern human-like life histories, social behavior, and locomotion (e.g., Foley & Gamble, 2009; Pradhan et al., 2012; Shea, 2017; Sterelny, 2021).

The importance of the Oldowan-early Acheulean transition may nonetheless be overemphasized, as much of the technological know-how apparent in early Acheulean assemblages is also present in Oldowan technology pre-transition (consider the presence of bifaces in later Oldowan assemblages; Leakey, 1971; Shea, 2010; see also Duke et al., 2021) and the co-occurrence of Acheulean and Oldowan technology post-'transition' (Lepre et al., 2011; Semaw et al., 2009, 2020). The appearance of early Acheulean forms in spandrels experiments (Moore & Perston, 2016; and debatably in this study, with the protobiface candidate) puts the presence of intentional shaping into serious doubt, while similarly implying that least-effort pathways are enough to generate early Acheulean artifacts.

Problematically, much of what we understand – from an experimental standpoint – about the Oldowan-early Acheulean transition is based on comparisons of Oldowan flaking with later representatives of Mode 2 knapping. Instead of early Acheulean forms, experiments often use late Acheulean handaxes modelled after the site of Boxgrove (Stout et al., 2014; Pargeter et al., 2019, 2020) or even foliate bifaces from the Middle Stone Age/Middle Paleolithic (Muller & Clarkson, 2016). Potentially, the early Acheulean-late Acheulean transition represents a more significant shift in hominin technological behaviors and cognitive abilities (Foley & Lahr, 2003), which would suggest that comparisons between the Oldowan and the later examples of Mode 2 exaggerate the differences presumed for the Oldowan and early Acheulean. Here, it is important to disavow ourselves of the assumption that handaxes are indeed a unifying feature for a single, unified Acheulean technological concept (cf. Lycett & Gowlett).

Indeed, it is our hypothesis that the hominins that produced late Acheulean technology were the first ones to also exhibit human-like cumulative culture of know-how (see van Schaik et al., 2019). Late Acheulean hominins were distinct from those who first produced handaxes in the early Acheulean (e.g., at Boxgrove, the hominins would have been Middle Pleistocene or Neandertal lineage hominins; Lockey et al., 2022), having much larger brains that were approximately in the same range as modern humans (see Foley & Gamble, 2009; Hublin & Changeaux, 2022; Stout & Chaminade, 2007). In the late Acheulean, hominins were engaged in a much wider range of technological behaviors and know-how expression, including such innovations as platform preparation, tranchet removals, and manufacture of organic hammers from antler (Nowell & White, 2010; Stout et al., 2014; though consider the possibility that organic hammers were present before the late Acheulean: Bandini et al., 2021b; Clément, 2021; Luncz et al., 2022b). The diversification in toolmaking know-how implies that the early-late Acheulean transition represents a critical period for the evolution and emergence of cognitive and cultural mechanisms that are more along the lines of those in living modern humans. This conclusion is further supported by the application of phylogenetic inference. Growing evidence has demonstrated the cognitive and cultural complexity of Neandertals (e.g., Hoffecker, 2018; Hoffmann et al., 2018a, 2018b; Sykes, 2021), suggesting that certain cognitive and cultural capacities must have been present in the last common ancestor of Neandertals and modern humans. The estimated dates for the divergence of the two taxa just happens to coincide with the appearance of many of the behavioral innovations that typify the late Acheulean industry as well as some of the earliest potential signs of symbolic expression by hominins (Jelinek, 2001; Nowell & White, 2010; Stout et al., 2014; Wynn et al., 2019).

And in the meantime, Neandertals, modern humans, and other late Acheulean handaxe-makers also were producers of Mode 1 and early Acheulean forms, indicating the probably co-existence of minimal cultural processes with cumulative culture of know-how, which is further exemplified by the re-innovation of Oldowan-like technology by modern human test participants (e.g., Bulut et al., 2022; Mithen, 1996; Pacheco et al., 2012; Snyder et al., 2022).

Research Outlook

Experimental archaeology, at least in the current regard of controlled experimentation and investigation of evolutionary processes related to cognition, culture, and biology, is a nascent and burgeoning discipline (e.g., Eren et al., 2016). There is plenty more empirical research that needs to be done, also in the direct study of the true archaeological record.

In particular, effort should be made to carry longitudinal studies of – above all – naïve individuals (longitudinal studies are rare and relatively short testing times in other studies may be poorly informative: see Pargeter et al., 2019, 2020, 2021; Stout et al., 2009). Additional longitudinal studies would help us better understand the skill acquisition process in a way unachievable through other methodologies. In our case, we were unfortunately unable to pursue the remainder of our intended testing program – an additional planned eighteen hours – due to the advent of the COVID-19 global health crisis. Future research would ideally make up for this, by pursuing longer-term evaluation of naïve knapper capabilities.

We would predict that, when given enough time to practice early knapping techniques (see Pargeter et al., 2021), (totally or technique-) naïve knappers would be capable of eventually reaching the kind of 'expert'-level performance that is known from early Oldowan sites (e.g., Stout & Semaw, 2006; contra predictions of Eren et al., 2020). Indeed, with upwards of 300 hours of testing or practice time and suitable raw materials (i.e., initial blank forms), early Acheulean core forms and reduction patterns should also be possible, noting the 'foreshadowing' offered by the protobiface candidate from this study and also the results of spandrels experiments (Moore & Perston, 2016). Whereas late Acheulean forms are not a strong guarantee to appear (requiring 200 hours or more when novices are given demonstrations and guidance; Pargeter et al., 2019), simple versions of Levallois would even be possible, essentially as a development upon the high frequency of centripetal knapping in this study (with an intermediate stage of bifacial hierarchical knapping as known from later Oldowan sites, see de la Torre, 2011; de la Torre et al., 2003; E. Moos, personal comm.).

Although the results of this study should still serve as viable proof-of-principle for the general notion that early knapping techniques and Oldowan artifacts can be re-innovated individually, good scientific praxis recommends that the results be replicated in further populations and under modified testing conditions. One aspect of this (to overcome the overreliance on WEIRD populations: e.g., Bandini et al., 2022; Henrich et al., 2010; Killin & Pain, 2022; Snyder et al., 2022; Wynn, 1985) is corroborate the results of this study with results from tests using baseline conditions in other populations around the world. Another avenue of research is to investigate the knapping abilities of children, to further elaborate on the developmental component of skill acquisition (cf. Bandini et al., 2022; Ferguson, 2003; Neldner et al., 2020; Reindl et al., 2016; Sternke & Sørensen, 2007). Tests with human children would also benefit from being performed cross-culturally (cf. Neldner et al., 2020; Reindl et al., 2016). Finally, the island test conditions could be modified to include, e.g., different raw materials, a new motivational task, and in a different environment (such as a naturalistic riverbed setting; D. Parris, personal comm.). We would predict that re-innovation should occur in all these cases, and – as would still fit the minimal culture model – changes in, e.g., participant demographics or knappable raw materials would lead to differences in the frequency of behavioral re-innovations and the artifact types (e.g., Toth, 1985).

Next, we recommend testing of other behaviors and technological components that we were not able to cover in the scope of our study and that have generally been tested in copying social learning paradigms (as with early knapping techniques, this may have been erroneous). Some examples of testable elements include core rotation (e.g., Kimbel et al., 1996; Stout et al., 2019), long-distance procurement of raw materials (e.g., Reeves et al., 2021), raw material selectivity (e.g., Duke & Pargeter, 2015), flake modification, side-scrapers, and other flake-blank technologies (appeared somewhat in our study, but still requires additional analysis and further testing; Gabrić et al., 2021), tool use tasks that are directly related to the Oldowan like butchering and wood processing, nut cracking and other percussive tool use behaviors (looking at adult modern humans and species that have not yet been tested for re-innovation capacities), hafting (see Hayden, 2015 for hafting of simple Mode 1 flakes by Australian Aborigines), and – last but not least – know-how related to the Acheulean and later technocomplexes from prehistory.

No part of this research program is to be conducted in isolation. Instead, novice knapping experiments pursuing the aforementioned research topics should be triangulated with empirical data gained from other sources and methodologies. Naturally, the archaeological record is an important component of this; more cross-site comparisons are needed, as well as more concentrated efforts on evaluating variability in the record, in order to under changes and stability in technology at both the micro- and macro-level. As has been emphasized throughout this dissertation, cross-species comparisons of know-how related to both toolmaking and tool use are highly consequential for interpreting both the past and present of hominin and human toolmaking (Bandini et al., 2022; Henrich et al., 2010; Killin & Pain, 2022; Tennie et al., 2020). Future work is needed to conduct more direct comparisons between unintentional and intentional flaking of apes and novice humans, which might help to inform us about pre-Oldowan tool use and the origins of toolmaking, while also helping to uncover traces of accidental discovery events within the Oldowan and later times (Bandini et al. 2021c, 2022; Motes-Rodrigo et al. 2022; Proffitt et al. 2016, 2018, 2022a; Stout et al., 2009). Finally, computational modelling (e.g., agent-based models) can be a useful strategy for studying larger scale processes that cannot readily be examined in controlled experiments or which have no living models we can use (e.g., looking at how mild cross-species distinctions in cognition can influence cultural outcomes, as e.g., bonobo to chimpanzee comparisons might not sufficiently help to elucidate such archaeological conundrums core rotation at Gona sites, e.g., Stout et al., 2019).

The results of this study (Snyder et al., 2022) have not just provided valuable insights into cognition and culture of some of the earliest toolmakers in the hominin lineage, but they have also demonstrated the viability of our methodological approach, thereby opening up many new possibilities for investigation. As the field of experimental cognitive archaeology moves forward, the onus is on researchers to examine questions of prehistoric culture and cognition from as many directions as possible and with as many methodologies as are within reason.

References

A

Acerbi, A., Van Leeuwen, E. J., Haun, D., & Tennie, C. (2016). Conformity cannot be identified based on population-level signatures. *Scientific Reports*, *6*(1), 1-9.

Acerbi, A., Snyder, W.D., & Tennie, C. (2022). The method of exclusion (still) cannot identify mechanisms of cultural inheritance. *Scientific Reports, 12*, 21680.

Allritz, M., Tennie, C., & Call, J. (2013). Food washing and placer mining in captive great apes. *Primates*, *54*(4), 361-370.

Ambrose, S. H. (2001). Paleolithic technology and human evolution. *Science*, *291*(5509), 1748-1753.

Apel, J. (2001). *Daggers, Knowledge & Power* (Doctoral dissertation, Acta Universitatis Upsaliensis).

Arbilly, M., & Laland, K. N. (2014). The local enhancement conundrum: in search of the adaptive value of a social learning mechanism. *Theoretical Population Biology*, *91*, 50-57.

B

Baber, C., & Janulis, K. (2020). Purposeful tool use in early lithic technologies. *Adaptive Behavior*, *29*(2), 169-180.

Backwell, L., & d'Errico, F. (2008). Early hominid bone tools from Drimolen, South Africa. *Journal of Archaeological Science*, *35*(11), 2880-2894.

Bamforth, D. B., & Finlay, N. (2008). Introduction: archaeological approaches to lithic production skill and craft learning. *Journal of Archaeological Method and Theory*, *15*(1), 1-27.

Bandini, E., & Tennie, C. (2017). Spontaneous reoccurrence of "scooping", a wild tool-use behaviour, in naïve chimpanzees. *PeerJ*, *5*, e3814.

Bandini, E., & Tennie, C. (2018). Naive, captive long-tailed macaques (*Macaca fascicularis fascicularis*) fail to individually and socially learn pound-hammering, a tool-use behaviour. *Royal Society Open Science*, *5*(5), 171826.

Bandini, E., & Tennie, C. (2019). Individual acquisition of "stick pounding" behavior by naïve chimpanzees. *American Journal of Primatology*, e22987.

Bandini, E., & Tennie, C. (2020). Exploring the role of individual learning in animal tool-use. *PeerJ*, *8*, e9877.

Bandini, E., Motes-Rodrigo, A., Steele, M. P., Rutz, C., & Tennie, C. (2020). Examining the mechanisms underlying the acquisition of animal tool behaviour. *Biological Letters*, *16*, 20200122.

Bandini, E., Reeves, J. S., Snyder, W. D., & Tennie, C. (2021a). Clarifying misconceptions of the zone of latent solutions hypothesis: a response to Haidle and Schlaudt. *Biological Theory*, *16*(2), 76-82.

Bandini, E., Grossmann, J., Funk, M., Albiach-Serrano, A., & Tennie, C. (2021b). Naïve orangutans (*Pongo abelii* and *Pongo pygmaeus*) individually acquire nut-cracking using hammer tools. *American Journal of Primatology*, *83*(9), e23304.

Bandini, E., Motes-Rodrigo, A., Archer, W., Minchin, T., Axelsen, H., Hernandez-Aguilar, R. A., McPherron, S. P., & Tennie, C. (2021c). Naïve, unenculturated chimpanzees fail to make and use flaked stone tools. *Open Research Europe*, *1*, 20.

Bandini, E., Harrison, R. A., & Motes-Rodrigo, A. (2022). Examining the suitability of extant primates as models of hominin stone tool culture. *Humanities and Social Sciences Communications*, *9*(1), 1-18.

Bargalló, A., & Mosquera, M. (2014). Can hand laterality be identified through lithic technology? *Laterality: Asymmetries of Body, Brain and Cognition*, *19*(1), 37–63.

Bargalló, A., Mosquera, M., & Lorenzo, C. (2018). Identifying handedness at knapping; an analysis of the scatter pattern of lithic remains. *Archaeological and Anthropological Sciences*, *10*(3), 587-598.

Barrett, B. J. (2019). Equifinality in empirical studies of cultural transmission. *Behavioural Processes*, *161*, 129-138.

Barsky, D. (2009). An overview of some African and Eurasian Oldowan sites: evaluation of hominin cognition levels, technological advancement and adaptive skills. *Interdisciplinary Approaches to the Oldowan*, 39-47.

Bayani, K. Y. T., Natraj, N., Khresdish, N., Pargeter, J., Stout, D., & Wheaton, L. A. (2021). Emergence of perceptuomotor relationships during paleolithic stone toolmaking learning: Intersections of observation and practice. *Communications Biology*, *4*(1), 1278.

Beck, B. B. (1974). Baboons, chimpanzees, and tools. *Journal of Human Evolution*, *3*(6), 509-516.

Bernstein-Kurtycz, L. M., Hopper, L. M., Ross, S. R., & Tennie, C. (2020). Zoo-housed chimpanzees can spontaneously use tool sets but perseverate on previously successful tool-use methods. *Animal Behavior and Cognition*, *7*(3), 288-309.

Beyin, A., Chauhan, P. R., & Nassr, A. (2017). New discovery of Acheulean occupation in the Red Sea coastal region of the Sudan. *Evolutionary Anthropology: Issues, News, and Reviews*, *26*(6), 255-257.

Binford, L. 1972. *An Archaeological Perspective*. Seminar Press.

Binford, L. R., & O'Connell, J. F. (1984). An Alyawara day: The stone quarry. *Journal of Anthropological Research*, *40*(3), 406-432.

Biro, D., Inoue-Nakamura, N., Tonooka, R., Yamakoshi, G., Sousa, C., & Matsuzawa, T. (2003). Cultural innovation and transmission of tool use in wild chimpanzees: evidence from field experiments. *Animal Cognition*, *6*(4), 213-223.

Bisson, M. S. (2001). Interview with a Neanderthal: An Experimental Approach for Reconstructing Scraper Production Rules, and their Implications for Imposed Form in Middle Palaeolithic Tools. *Cambridge Archaeological Journal*, *11*(2), 165–184.

Boesch, C., & Boesch-Achermann, H. (2000). *The Chimpanzees of the Taï Forest: Behavioural Ecology and Evolution*. Oxford University Press, USA.

Boesch, C., & Tomasello, M. (1998). Chimpanzee and human cultures. *Current Anthropology*, *39*(5), 591-614.

Boyd, R. (2018). *A Different Kind of Animal: How Culture Transformed Our Species*. Princeton University Press.

Boyd, R., & Richerson, P. J. (1988). *Culture and the Evolutionary Process*. University of Chicago Press.

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

Boyd, R., & Richerson, P. J. (1996, January). Why culture is common, but cultural evolution is rare. In *Proceedings of the British Academy* (Vol. 88, pp. 77-94). Oxford University Press.

Braun, D. R., Tactikos, J. C., Ferraro, J. V., & Harris, J. W. K. (2005). Flake recovery rates and inferences of Oldowan hominin behavior: a response to Kimura 1999, 2002. *Journal of Human Evolution*, *48*(5), 525-531.

Braun, D. R., Plummer, T., Ferraro, J. V., Ditchfield, P., & Bishop, L. C. (2009). Raw material quality and Oldowan hominin toolstone preferences: evidence from Kanjera South, Kenya. *Journal of Archaeological Science*, *36*(7), 1605-1614.

Braun, D. R., Aldeias, V., Archer, W., Arrowsmith, J. R., Baraki, N., Campisano, C. J., Deino, A. L., DiMaggio, E. N., Dupont-Nivet, G., Engda, B., Feary, D. A., Garello, D. I., Kerfelew, Z., McPherron, S. P., Patterson, D. B., Reeves, J. S., Thompson, J. C., & Reed, K. E. (2019). Earliest known Oldowan artifacts at >2.58 Ma from Ledi-Geraru, Ethiopia, highlight early technological diversity. *Proceedings of the National Academy of Sciences*, 201820177.

Bril, B., Rein, R., Nonaka, T., Wenban-Smith, F., & Dietrich, G. (2010). The role of expertise in tool use: skill differences in functional action adaptations to task constraints. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(4), 825.

Bril, B., Parry, R., & Dietrich, G. (2015). How similar are nut-cracking and stone-flaking? A functional approach to percussive technology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1682), 20140355.

Brown, F. H., & Gathogo, P. N. (2002). Stratigraphic Relation between Lokalalei 1A and Lokalalei 2C, Pliocene Archaeological Sites in West Turkana, Kenya. *Journal of Archaeological Science*, *29*(7), 699-702.

Bruner, E., Preuss, T. M., Chen, X., & Rilling, J. K. (2017). Evidence for expansion of the precuneus in human evolution. *Brain Structure and Function*, *222*(2), 1053-1060.

Bulut, H., Taşkıran, H., Özçelik, K., & Karahan, G. (2022). Lower and Middle Palaeolithic evidence from the North Aegean coastline of Çanakkale, Turkey. *Antiquity*, 1-8.

Buskell, A., & Tennie, C. (in press). Mere recurrence and cumulative culture at the margins. *The British Journal for the Philosophy of Science*.

Byrne, F., Proffitt, T., Arroyo, A., & de la Torre, I. (2016). A comparative analysis of bipolar and freehand experimental knapping products from Olduvai Gorge, Tanzania. *Quaternary International*, *424*, 58-68.

Byrne, R. W. (2007). Culture in great apes: using intricate complexity in feeding skills to trace the evolutionary origin of human technical prowess. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1480), 577-585.

Byrne, R. W. (2009). Animal imitation. *Current Biology*, *19*(3), R111-R114.

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

Byrne, R. W., Cartmill, E., Genty, E., Graham, K. E., Hobaiter, C., & Tanner, J. (2017). Great ape gestures: intentional communication with a rich set of innate signals. *Animal Cognition*, *20*(4), 755-769.

C

Caldwell, C. A., & Millen, A. E. (2009). Social learning mechanisms and cumulative cultural evolution: is imitation necessary?. *Psychological Science*, *20*(12), 1478-1483.

Call, J., & Tomasello, M. (1996). The effect of humans on the cognitive development of apes. *Reaching into Thought: The Minds of the Great Apes*, 371-403.

Callahan, E. (1979). The basics of biface knapping in the eastern fluted point tradition: a manual for flintknappers and lithic analysts. *Archaeology of Eastern North America*, *7*, 1–180.

Carbonell, E., Bermúdez de Castro, J. M., Arsuaga, J. L., Diez, J. C., Rosas, A., Cuenca-Bescós, G., ... & Rodríguez, X. P. (1995). Lower Pleistocene hominids and artifacts from Atapuerca-TD6 (Spain). *Science*, *269*(5225), 826-830.

Carbonell, E., Garcı, M., Mallol, C., Mosquera, M., Ollé, A., Sahnouni, M., ... & Vergs, J. M. (1999). The TD6 level lithic industry from Gran Dolina, Atapuerca (Burgos, Spain): production and use. *Journal of Human Evolution*, *37*(3-4), 653-693.

Caruana, M. V., d'Errico, F., & Backwell, L. (2013). Early hominin social learning strategies underlying the use and production of bone and stone tools. In C. Sanz, J. Call, & C. Boesch (Eds.), *Tool Use in Animals* (pp. 242–285). Cambridge University Press.

Carvalho, S., & McGrew, W. C. (2012). The origins of the Oldowan: Why chimpanzees (*Pan troglodytes*) still are good models for technological evolution in Africa. *Stone Tools and Fossil Bones: Debates in the Archaeology of Human Origins*, 201-221

Carvalho, S., Cunha, E., Sousa, C., & Matsuzawa, T. (2008). Chaînes opératoires and resource-exploitation strategies in chimpanzee (*Pan troglodytes*) nut cracking. *Journal of Human Evolution*, *55*(1), 148-163.

Carvalho, S., Matsuzawa, T., McGrew, W. C., Sanz, C., Call, J., & Boesch, C. (2013). From pounding to knapping: how chimpanzees can help us to model hominin lithics. *Tool Use in Animals: Cognition and Ecology*, 225-241.

Cataldo, D. M., Migliano, A. B., & Vinicius, L. (2018). Speech, stone tool-making and the evolution of language. *PLoS ONE*, *13*(1), e0191071.

Cerasoni, J. N. (2021). Vectorial application for the illustration of archaeological lithic artefacts using the "Stone Tools Illustrations with Vector Art" (STIVA) Method. *PLoS ONE*, *16*, 0251466.

Childe, V. G. (1929). *The Danube in Prehistory*. Oxford: Clarendon Press.

Clark, G. (1963). *World Prehistory: A New Outline*. Cambridge: Cambridge University Press.

Clarkson, C. (2017). Teaching Complex flint knapping strategies in the classroom using "potato knapping". *Lithic Technology*, *42*(4), 155-160.

Clay, Z., & Tennie, C. (2018). Is Overimitation a Uniquely Human Phenomenon? Insights From Human Children as Compared to Bonobos. *Child Development*, *89*(5), 1535–1544.

Clément, S. (2021). Soft-Hammer Percussion During the Acheulean: Barking Up the Wrong Tree of Technical Change?. *Journal of Paleolithic Archaeology*, *5*(1), 1-30.

Collias, E. C., & Collias, N. E. (1964). The Development of Nest-Building Behavior in a Weaverbird. *The Auk*, *81*(1), 42–52.

Cooper, Z., & Bowdler, S. (1998). Flaked glass tools from the Andaman Islands and Australia. *Asian Perspectives*, 74-83.

Corbey, R. (2020). Baldwin effects in early stone tools. *Evolutionary Anthropology: Issues, News, and Reviews*, *29*(5), 237–244.

Corbey, R., Jagich, A., Vaesen, K., & Collard, M. (2016). The acheulean handaxe: More like a bird's song than a beatles' tune? *Evolutionary Anthropology: Issues, News, and Reviews*, *25*(1), 6–19.

Cotterell, B., Kamminga, J., & Dickson, F. P. (1985). The essential mechanics of conchoidal flaking. *International Journal of Fracture*, *29*(4), 205-221.

Cotterell, B., & Kamminga, J. (1987). The formation of flakes. *American Antiquity*, *52*, 675–708.

Cueva-Temprana, A., Lombao, D., Soto, M., Itambu, M., Bushozi, P., Boivin, N., ... & Mercader, J. (2022). Oldowan Technology Amid Shifting Environments∼ 2.03–1.83 Million Years Ago. *Frontiers in Ecology and Evolution*, 122.

D

Danchin, E., Nöbel, S., Pocheville, A., Dagaeff, A.-C., Demay, L., Alphand, M., Ranty-Roby, S., van Renssen, L., Monier, M., Gazagne, E., Allain, M., & Isabel, G. (2018). Cultural flies: Conformist social learning in fruitflies predicts long-lasting mate-choice traditions. *Science, 362*(6418), 1025-1030.

Davidson, I. (2002). The finished artefact fallacy: Acheulean hand-axes and language origins. In *The Transition to Language*. Oxford University Press.

Davidson, I., & Noble, W. (1993). Tools and language in human evolution. *Tools, Language and Cognition in Human Evolution*, *363*(388), 213-229.

Davidson, I., & McGrew, W. C. (2005). Stone tools and the uniqueness of human culture. *Journal of the Royal Anthropological Institute*, *11*(4), 793-817.

Davis, S. J., Vale, G. L., Schapiro, S. J., Lambeth, S. P., & Whiten, A. (2016). Foundations of cumulative culture in apes: improved foraging efficiency through relinquishing and combining witnessed behaviours in chimpanzees (*Pan troglodytes*). *Scientific Reports*, *6*(1), 1-12.

Dean, L. G., Kendal, R. L., Schapiro, S. J., Thierry, B., & Laland, K. N. (2012). Identification of the social and cognitive processes underlying human cumulative culture. *Science*, *335*(6072), 1114-1118.

Dean, L. G., Vale, G. L., Laland, K. N., Flynn, E., & Kendal, R. L. (2013). Human cumulative culture: a comparative perspective. *Biological Reviews*, *89*(2), 284-301.

De Beaune, S. (2004). The invention of technology: prehistory and cognition. *Current Anthropology*, *45*(2), 139-162.

Delagnes, A., & Roche, H. (2005). Late Pliocene hominid knapping skills: the case of Lokalalei 2C, West Turkana, Kenya. *Journal of Human Evolution*, *48*(5), 435-472.

de la Torre, I. (2004). Omo revisited: evaluating the technological skills of Pliocene hominids. *Current Anthropology*, *45*(4), 439-465.

de la Torre, I. (2011). The Early Stone Age lithic assemblages of Gadeb (Ethiopia) and the Developed Oldowan/early Acheulean in East Africa. *Journal of Human Evolution*, *60*(6), 768–812.

de la Torre, I. (2016). The origins of the Acheulean: past and present perspectives on a major transition in human evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1698), 20150245.

de la Torre, I. (2019). Searching for the emergence of stone tool making in eastern Africa. *Proceedings of the National Academy of Sciences*, *116*(24), 11567-11569.

de la Torre, I., & Mora, R. (2005). Technological Strategies in the Lower Pleistocene at Olduvai Beds I & II. *ERAUL*, *112*.

de la Torre, I., & Mora, R. (2013). The transition to the Acheulean in East Africa: an assessment of paradigms and evidence from Olduvai Gorge (Tanzania). *Journal of Archaeological Method and Theory*, *21*(4), 781-823.

de la Torre, I., Mora, R., Domínguez-Rodrigo, M., de Luque, L., & Alcalá, L. (2003). The Oldowan industry of Peninj and its bearing on the reconstruction of the technological skills of Lower Pleistocene hominids. *Journal of Human Evolution*, *44*(2), 203-224.

de Lumley, H. & Beyene, Y. (2004). *Les sites préhistoriques de la région de Fejej, Sud-Omo, Ethiopie, dans leur contexte stratigraphique et paléontologique*. Paris: Association pour la Diffusion de la Pensée Française.

d'Errico, F., & Backwell, L. R. (2003). Possible evidence of bone tool shaping by Swartkrans early hominids. *Journal of Archaeological Science*, *30*(12), 1559-1576.

Dibble, H. L., & Whittaker, J. C. (1981). New experimental evidence on the relation between percussion flaking and flake variation. *Journal of Archaeological Science*, *8*(3), 283-296.

Dibble, H. L., & Pelcin, A. (1995). The effect of hammer mass and velocity on flake mass. *Journal of Archaeological Science*, *22*(3), 429-439.

Dibble, H. L., & 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*(9), 1945-1954.

Dindo, M., Thierry, B., & Whiten, A. (2007). Social diffusion of novel foraging methods in brown capuchin monkeys (*Cebus apella*). *Proceedings of the Royal Society B: Biological Sciences*, *275*(1631), 187-193.

Dogandžić, T., Abdolazadeh, A., Leader, G., Li, L., McPherron, S. P., Tennie, C., & Dibble, H. L. (2020). The results of lithic experiments performed on glass cores are applicable to other raw materials. *Archaeological and Anthropological Sciences*, *12*(2), 44.

Domínguez-Rodrigo, M. (2009). Are all Oldowan Sites Palimpsests? If so, what can they tell us about Hominid Carnivory?. In *Interdisciplinary approaches to the Oldowan* (pp. 129-147). Springer, Dordrecht.

Domínguez-Rodrigo, M., & Alcalá, L. (2016). 3.3-million-year-old stone tools and butchery traces? More evidence needed. *PaleoAnthropology*, *2016*, 46-53.

Donald, M. (1993). Hominid enculturation and cognitive evolution. *Archaeological Review from Cambridge*, *12*(2), 5-24.

Duke, H., & Pargeter, J. (2015). Weaving simple solutions to complex problems: An experimental study of skill in bipolar cobble-splitting. *Lithic Technology*, *40*(4), 349–365.

Duke, H., Feibel, C., & Harmand, S. (2021). Before the Acheulean: the emergence of bifacial shaping at Kokiselei 6 (1.8 Ma), West Turkana, Kenya. *Journal of Human Evolution*, *159*, 103061.

Dumoncel, J., Subsol, G., Durrleman, S., Bertrand, A., de Jager, E., Oettlé, A. C., ... & Beaudet, A. (2021). Are endocasts reliable proxies for brains? A 3D quantitative comparison of the extant human brain and endocast. *Journal of Anatomy*, *238*(2), 480-488.

Dusseldorp, G. L., & Lombard, M. (2021). Constraining the likely technological niches of late Middle Pleistocene hominins with Homo naledi as case study. *Journal of Archaeological Method and Theory*, *28*(1), 11-52.

E

Eerkens, J. W., & Lipo, C. P. (2005). Cultural transmission, copying errors, and the generation of variation in material culture and the archaeological record. *Journal of Anthropological Archaeology*, *24*(4), 316-334.

Eren, M. I., Lycett, S. J., Patten, R. J., Buchanan, B., Pargeter, J., & O'Brien, M. J. (2016). Test, model, and method validation: The role of experimental stone artifact replication in hypothesis-driven archaeology. *Ethnoarchaeology, 8*(2), 103–136.

Eren, M. I., Lycett, S. J., & Tomonaga, M. (2020). Underestimating Kanzi? Exploring Kanzi-Oldowan comparisons in light of recent human stone tool replication. *Evolutionary Anthropology: Issues, News, and Reviews*, *29*(6), 310-316.

F

Falótico, T., Proffitt, T., Ottoni, E. B., Staff, R. A., & Haslam, M. (2019). Three thousand years of wild capuchin stone tool use. *Nature Ecology & Evolution*, *3*(7), 1034-1038.

Farmer, G. L., Holland, H. D., & Turekian, K. K. (2003). Continental basaltic rocks. *The Crust*, *3*, 85-121.

Ferguson, J. R. (2003). An experimental test of the conservation of raw material in flintknapping skill acquisition. *Lithic Technology*, *28*(2), 113–131.

Flenniken, J. J. (1984). The past, present, and future of flintknapping: an anthropological perspective. *Annual Review of Anthropology*, 187-203.

Foley, R. (1987). Hominid species and stone-tool assemblages: how are they related?. *Antiquity*, *61*(233), 380-392.

Foley, R. A. (2016). Mosaic evolution and the pattern of transitions in the hominin lineage. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1698), 20150244.

Foley, R. A., & Elton, S. (1998). Time and energy: the ecological context for the evolution of bipedalism. In *Primate Locomotion* (pp. 419-433). Springer, Boston, MA.

Foley, R., & Gamble, C. (2009). The ecology of social transitions in human evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1533), 3267-3279.

Foley, R., & Lahr, M. M. (2003). On stony ground: lithic technology, human evolution, and the emergence of culture. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, *12*(3), 109-122.

Forss, S. I. F., Motes-Rodrigo, A., Hrubesch, C., & Tennie, C. (2019). Differences in novel food response between *Pongo* and *Pan*. *American Journal of Primatology*, *81*(1), e22945.

Fragaszy, D. M., Biro, D., Eshchar, Y., Humle, T., Izar, P., Resende, B., & Visalberghi, E. (2013). The fourth dimension of tool use: temporally enduring artefacts aid primates learning to use tools. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1630), 20120410.

Furlong, E. E., Boose, K. J., & Boysen, S. T. (2008). Raking it in: the impact of enculturation on chimpanzee tool use. *Animal Cognition*, *11*(1), 83-97.

G

Gabora, L., & Steel, M. (2020). Modeling a cognitive transition at the origin of cultural evolution using autocatalytic networks. *Cognitive Science*, *44*(9), e12878.

Gabrić, P., Banda, M., & Karavanić, I. (2021). Cognitive performance and specific aspects of language processing are associated with Oldowan-like chert flaking and retouch. *Research Square*. https://doi.org/10.21203/rs.3.rs-1212115/v1

Gabunia, L., Vekua, A., Lordkipanidze, D., Swisher III, C. C., Ferring, R., Justus, A., ... & Mouskhelishvili, A. (2000). Earliest Pleistocene hominid cranial remains from Dmanisi, Republic of Georgia: taxonomy, geological setting, and age. *Science*, *288*(5468), 1019-1025.

Gallotti, R. (2018). Before the Acheulean in East Africa: An overview of the Oldowan lithic assemblages. *The Emergence of the Acheulean in East Africa and Beyond*, 13-32.

Gallotti, R., & Mussi, M. (2015). The unknown Oldowan: ~1.7-million-year-old standardized obsidian small tools from Garba IV, Melka Kunture, Ethiopia. *PLoS ONE*, *10*(12), e0145101.

Gärdenfors, P., & Högberg, A. (2017). The archaeology of teaching and the evolution of Homo docens. *Current Anthropology*, *58*(2), 188-208.

Geribàs, N., Mosquera, M., & Vergès, J. M. (2010). What novice knappers have to learn to become expert stone toolmakers. *Journal of Archaeological Science*, *37*(11), 2857-2870.

Goldman-Neuman, T., & Hovers, E. (2011). Raw material selectivity in late Pliocene Oldowan sites in the Makaamitalu Basin, Hadar, Ethiopia. *Journal of Human Evolution*, *62*(3), 353-366.

Gowlett, J. A. (1996). Mental abilities of early Homo: elements of constraint and choice in rule systems. *Modelling the Early Human Mind*, 191-215.

Gruber, T., & Clay, Z. (2016). A comparison between bonobos and chimpanzees: A review and update. *Evolutionary Anthropology: Issues, News, and Reviews*, *25*(5), 239-252.

Gürbüz, R. B., & Lycett, S. J. (2021). Did the use of bone flakes precede the use of knapped stone flakes in hominin meat processing and could this be detectable archaeologically?. *Journal of Anthropological Archaeology*, *62*, 101305.

Gumert, M. D., Tan, A. W. Y., Luncz, L. V., Chua, C. T., Kulik, L., Switzer, A. D., ... & Malaivijitnond, S. (2019). Prevalence of tool behaviour is associated with pelage phenotype in intraspecific hybrid long-tailed macaques (*Macaca fascicularis aurea× M. f. fascicularis*). *Behaviour*, *156*(11), 1083-1125.

Gunz, P., Neubauer, S., Maureille, B., & Hublin, J. J. (2010). Brain development after birth differs between Neanderthals and modern humans. *Current Biology*, *20*(21), R921-R922.

Gunz, P., Neubauer, S., Golovanova, L., Doronichev, V., Maureille, B., & Hublin, J. J. (2012). A uniquely modern human pattern of endocranial development. Insights from a new cranial reconstruction of the Neandertal newborn from Mezmaiskaya. *Journal of Human Evolution*, *62*(2), 300-313.

Gunz, P., Neubauer, S., Falk, D., Tafforeau, P., Le Cabec, A., Smith, T. M., ... & Alemseged, Z. (2020). *Australopithecus afarensis* endocasts suggest ape-like brain organization and prolonged brain growth. *Science Advances*, *6*(14), eaaz4729.

Gurtov, A. N., & Eren, M. I. (2014). Lower Paleolithic bipolar reduction and hominin selection of quartz at Olduvai Gorge, Tanzania: What's the connection?. *Quaternary International*, *322*, 285-291.

H

Haidle, M. N. (2010). Working-memory capacity and the evolution of modern cognitive potential: implications from animal and early human tool use. *Current Anthropology*, *51*(S1), S149-S166.

Haidle, M. N., & Schlaudt, O. (2020). Where Does Cumulative Culture Begin? A Plea for a Sociologically Informed Perspective. *Biological Theory*, *15*(3), 161–174.

Hamilton, W. J., Buskirk, R. E., & Buskirk, W. H. (1975). Defensive stoning by baboons. *Nature*, *256*(5517), 488-489.

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

Harlacker, L. (2003). Knowledge and know-how in the Oldowan: an experimental approach. *Skilled Production and Social Reproduction*, 219.

Harmand, S. (2009). Variability in raw material selectivity at the Late Pliocene sites of Lokalalei, West Turkana, Kenya. In *Interdisciplinary Approaches to the Oldowan* (pp. 85-97). Springer, Dordrecht.

Harmand, S., Lewis, J. E., Feibel, C. S., Lepre, C. J., Prat, S., Lenoble, A., Boës, X., Quinn, R. L., Brenet, M., Arroyo, A., Taylor, N., Clément, S., Daver, G., Brugal, J.-P., Leakey, L., Mortlock, R. A., Wright, J. D., Lokorodi, S., Kirwa, C., … Roche, H. (2015). 3.3-million-year-old stone tools from Lomekwi 3, West Turkana, Kenya. *Nature, 521*(7552), 310–315.

Harris, J. W., & Isaac, G. (1976). The Karari industry: early Pleistocene archaeological evidence from the terrain east of Lake Turkana, Kenya. *Nature, 262*(5564), 102-107.

Haslam, M., Hernandez-Aguilar, A., Ling, V., Carvalho, S., De La Torre, I., DeStefano, A., ... & Warren, R. (2009). Primate archaeology. *Nature*, *460*(7253), 339-344.

Haslam, M., Hernandez-Aguilar, R. A., Proffitt, T., Arroyo, A., Falótico, T., Fragaszy, D., Gumert, M., Harris, J. W. K., Huffman, M. A., Kalan, A. K., Malaivijitnond, S., Matsuzawa, T., McGrew, W., Ottoni, E. B., Pascual-Garrido, A., Piel, A., Pruetz, J., Schuppli, C., Stewart, F., . . . Luncz, L. V. (2017). Primate archaeology evolves. *Nature Ecology &Amp; Evolution*, *1*(10), 1431–1437.

Hay, R. L. (1976). *Geology of the Olduvai Gorge: A Study of Sedimentation in a Semiarid Basin*. University of California Press.

Hayden, B. (2015). Insights into early lithic technologies from ethnography. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1682), 20140356.

Hecht, E. E., Gutman, D. A., Khreisheh, N., Taylor, S. V., Kilner, J., Faisal, A. A., Bradley, B. A., Chaminade, T., & Stout, D. (2015). Acquisition of Paleolithic toolmaking abilities involves structural remodeling to inferior frontoparietal regions. *Brain Structure and Function*, *220*(4), 2315–2331.

Henrich, J. (2016). *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton University Press.

Henrich, J., & Tennie, C. (2017). Cultural evolution in chimpanzees and humans. In *Chimpanzees and Human Evolution*. Harvard University Press.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83.

Hernández, V., Jorge-Villar, S., Capel Ferrón, C., Medianero, F. J., Ramos, J., Weniger, G. C., ... & Durán Valsero, J. J. (2012). Raman spectroscopy analysis of Palaeolithic industry from Guadalteba terrace river, Campillos (Guadalteba county, Southern of Iberian Peninsula). *Journal of Raman Spectroscopy*, *43*(11), 1651-1657.

Heyes, C. M. (1994). Social learning in animals: categories and mechanisms. *Biological Reviews*, *69*(2), 207-231.

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

Heyes, C. (2020). Culture. *Current Biology*, *30*(20).

Heyes, C. (2021). Imitation and culture: What gives?. *Mind & Language*.

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

Hiscock, P. (2015). Dynamics of knapping with bipolar techniques: modeling transitions and the implications of variability. *Lithic Technology*, *40*(4), 342-348.

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

Högberg, A. (2008). Playing with Flint: Tracing a Child's Imitation of Adult Work in a Lithic Assemblage. *Journal of Archaeological Method and Theory*, *15*(1), 112–131.

Hoffecker, J. F. (2018). The complexity of Neanderthal technology. *Proceedings of the National Academy of Sciences*, *115*(9), 1959-1961.

Hoffmann, D. L., Standish, C. D., García-Diez, M., Pettitt, P. B., Milton, J. A., Zilhão, J., ... & Pike, A. W. (2018). U-Th dating of carbonate crusts reveals Neandertal origin of Iberian cave art. *Science*, *359*(6378), 912-915.

Hoffmann, D. L., Angelucci, D. E., Villaverde, V., Zapata, J., & Zilhão, J. (2018). Symbolic use of marine shells and mineral pigments by Iberian Neandertals 115,000 years ago. *Science Advances*, *4*(2), eaar5255.

Hoppitt, W., & Laland, K. N. (2008). Social processes influencing learning in animals: a review of the evidence. *Advances in the Study of Behavior*, *38*, 105-165.

Horta, P., Bicho, N., & Cascalheira, J. (2022). Lithic bipolar methods as an adaptive strategy through space and time. *Journal of Archaeological Science: Reports*, *41*, 103263.

Hovers, E. (2009). Learning from mistakes: flaking accidents and knapping skills in the assemblage of AL 894 (Hadar, Ethiopia). *The Cutting Edge: New Approaches to the Archaeology of Human Origins*, 137-150.

Hovers, E. (2012). Invention, reinvention and innovation: the makings of Oldowan lithic technology. In *Developments in Quaternary Sciences* (Vol. 16, pp. 51-68). Elsevier.

Hublin, J. J., & Changeux, J. P. (2022). Paleoanthropology of cognition: an overview on hominin brain evolution. *Comptes Rendus. Biologies*, *345*(2), 1-19.

Hublin, J. J., Neubauer, S., & Gunz, P. (2015). Brain ontogeny and life history in Pleistocene hominins. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1663), 20140062.

Hublin, J. J., Ben-Ncer, A., Bailey, S. E., Freidline, S. E., Neubauer, S., Skinner, M. M., ... & Gunz, P. (2017). New fossils from Jebel Irhoud, Morocco and the pan-African origin of Homo sapiens. *Nature*, *546*(7657), 289-292.

Huffman, M. A., & Hirata, S. (2004). An experimental study of leaf swallowing in captive chimpanzees: insights into the origin of a self-medicative behavior and the role of social learning. *Primates*, *45*(2), 113-118.

Humle, T., & Matsuzawa, T. (2002). Ant-dipping among the chimpanzees of Bossou, Guinea, and some comparisons with other sites. *American Journal of Primatology: Official Journal of the American Society of Primatologists*, *58*(3), 133-148.

I

Inoue-Nakamura, N., & Matsuzawa, T. (1997). Development of stone tool use by wild chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *111*(2), 159.

Iovita, R., & McPherron, S. P. (2011). The handaxe reloaded: A morphometric reassessment of Acheulian and Middle Paleolithic handaxes. *Journal of Human Evolution*, *61*(1), 61-74.

Isaac, G. L. (1972). Early phases of human behaviour: models in Lower Palaeolithic archaeology. *Models in archaeology*, 167-199.

Isaac, G. L. (1976). Stages of cultural elaboration in the Pleistocene: possible archaeological indicators of the development of language capabilities. *Annals of the New York Academy of Sciences*, *280*(1), 275-288.

Isaac, G. L. (1977). *Olorgesailie: Archaeological Studies of a Middle Pleistocene Lake Basin in Kenya*. Chicago: The University of Chicago Press.

Isaac, G. L. (1984). The earliest archaeological traces. In J. D. Clark (Ed.), *The Cambridge History of Africa* (Vol. 1, pp. 157–247). Cambridge: Cambridge University Press.

Isaac, G., Harris, J. W. K., & Kroll, E. M. (1997). The stone artefact assemblages: a comparative study. In G. Isaac & A. B. B. Isaac (Eds.), *Koobi Fora Research Project Series* (Plio-Pleistocene Archaeology, Vol. 5, pp. 262–362). Oxford: Clarendon.

J

Jelinek, A. J. (1977). The Lower Paleolithic: Current Evidence and Interpretations. *Annual Review of Anthropology*, *6*, 11–32.

Jelínek, J. (2001). Some innovations and continuity in the behaviour of European Middle and Late Pleistocene Hominids. *Some Innovations and Continuity in the Behaviour of European Middle and Late Pleistocene Hominids*, 1000-1007.

JMP (2021). JMP 16.0.0. SAS Institute, Cary, NC URL https://www.jmp.com/de_de/home.html.

Johnson, L. L. (1978). A history of flint-knapping experimentation, 1838-1976. *Current Anthropology*, *19*(2), 337-372.

K

Karakostis, F. A., Haeufle, D., Anastopoulou, I., Moraitis, K., Hotz, G., Tourloukis, V., & Harvati, K. (2021). Biomechanics of the human thumb and the evolution of dexterity. *Current Biology*, *31*(6), 1317-1325.

Keller, W. D. (1981). The sedimentology of flint clay. *Journal of Sedimentary Research*, *51*(1), 233-244.

Kenward, B., Weir, A. A., Rutz, C., & Kacelnik, A. (2005). Tool manufacture by naive juvenile crows. *Nature*, *433*(7022), 121-121.

Key, A. J. M., & Lycett, S. J. (2014). Are bigger flakes always better? An experimental assessment of flake size variation on cutting efficiency and loading. *J. Archaeol. Sci*, *41*, 140–146.

Key, A., Proffitt, T., & de la Torre, I. (2020). Raw material optimization and stone tool engineering in the Early Stone Age of Olduvai Gorge (Tanzania). *Journal of the Royal Society Interface*, *17*(162), 20190377.

Khreisheh, N. N., Davies, D., & Bradley, B. A. (2013). Extending experimental control: the use of porcelain in flaked stone experimentation. *Advances in Archaeological Practice*, *1*(1), 38-46.

Killin, A., & Pain, R. (2021). Cognitive archaeology and the minimum necessary competence problem. *Biological Theory*, 1-15.

Killin, A., & Pain, R. (2022). How WEIRD is Cognitive Archaeology? Engaging with the Challenge of Cultural Variation and Sample Diversity. *Review of Philosophy and Psychology*, 1-25.

Kimbel, W. H., Walter, R. C., Johanson, D. C., Reed, K. E., Aronson, J. L., Assefa, Z., ... & Smith, P. E. (1996). Late Pliocene *Homo* and Oldowan Tools from the Hadar Formation (Kada Hadar Member), Ethiopia. *Journal of Human Evolution*, *31*(6), 549-561.

Kitahara-Frisch, J., & Norikoshi, K. (1982). Spontaneous sponge-making in captive chimpanzees. *Journal of Human Evolution*, *11*(1), 41-47.

Kivell, T. L. (2015). Evidence in hand: recent discoveries and the early evolution of human manual manipulation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1682), 20150105.

Klein, E. M., & Langmuir, C. H. (1989). Local versus global variations in ocean ridge basalt composition: A reply. *Journal of Geophysical Research: Solid Earth*, *94*, 4241-4252.

Koops, K., McGrew, W. C., & Matsuzawa, T. (2010). Do chimpanzees (*Pan troglodytes*) use cleavers and anvils to fracture *Treculia africana* fruits? Preliminary data on a new form of percussive technology. *Primates*, *51*(2), 175-178.

Koops, K., Schöning, C., Isaji, M., & Hashimoto, C. (2015). Cultural differences in ant-dipping tool length between neighbouring chimpanzee communities at Kalinzu, Uganda. *Scientific Reports*, *5*(1), 1-8.

Koops, K., Soumah, A. G., van Leeuwen, K. L., Camara, H. D., & Matsuzawa, T. (2022). Field experiments find no evidence that chimpanzee nut cracking can be independently innovated. *Nature Human Behaviour*, *6*(4), 487-494.

Kuman, K., Le Baron, J. C., & Gibbon, R. J. (2005). Earlier stone age archaeology of the Vhembe-Dongola National Park (South Africa) and vicinity. *Quaternary International*, *129*(1), 23-32.

Kuman, K., Sutton, M. B., Pickering, T. R., & Heaton, J. L. (2018). The Oldowan industry from Swartkrans cave, South Africa, and its relevance for the African Oldowan. *Journal of Human Evolution*, *123*, 52-69.

Kunze, J., Karakostis, F. A., Merker, S., Peresani, M., Hotz, G., Tourloukis, V., & Harvati, K. (2022). Entheseal patterns suggest habitual tool use in early hominins. *PaleoAnthropology*, *2022*(2).

Kurzban, R., & Barrett, H. C. (2012). Origins of cumulative culture. *Science*, *335*(6072), 1056-1057.

L

Langergraber, K. E., Boesch, C., Inoue, E., Inoue-Murayama, M., Mitani, J. C., Nishida, T., ... & Vigilant, L. (2010). Genetic and 'cultural' similarity in wild chimpanzees. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1704), 408-416.

Leader, G. M., Kuman, K., Gibbon, R. J., & Granger, D. E. (2018). Early Acheulean organised core knapping strategies ca. 1.3 Ma at Rietputs 15, Northern Cape Province, South Africa. *Quaternary International*, *480*, 16-28.

Leakey, M. D. (1971). *Olduvai Gorge: Volume 3, Excavations in Beds I and II, 1960-1963* (Vol. 3). Cambridge University Press.

Leakey, M. D. (1975). Cultural patterns in the Olduvai sequence. In K. W. Butzer, & G. L. Isaac (Eds.), *After the Australopithecines: Stratigraphy, Ecology and Culture Change in the Middle Pleistocene* (pp. 477-493). Paris: Mouton Publishers.

Lefebvre, L. (1986). Cultural diffusion of a novel food-finding behaviour in urban pigeons: An experimental field test. *Ethology*, *71*(4), 295-304.

Lewis, J. E., & Harmand, S. (2016). An earlier origin for stone tool making: implications for cognitive evolution and the transition to *Homo*. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1698), 20150233.

Lewis, H. M., & Laland, K. N. (2012). Transmission fidelity is the key to the build-up of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2171-2180.

Li, L., Lin, S. C., McPherron, S. P., Abdolahzadeh, A., Chan, A., Dogandžić, T., ... & Dibble, H. L. (2022). A Synthesis of the Dibble et al. Controlled Experiments into the Mechanics of Lithic Production. *Journal of Archaeological Method and Theory*, 1-42.

Lin, S. C., Rezek, Z., & Dibble, H. L. (2018). Experimental design and experimental inference in stone artifact archaeology. *Journal of Archaeological Method and Theory*, *25*(3), 663-688.

Lin, S.C., McNaughton, M., Clarkson, C., & Mackay, A. (2021). The use of a 3D milling machine in controlled flaking experiments: A case study of Levallois flake formation. *EXARC Experimental Archaeology Conference*.

Lockey, A. L., Rodríguez, L., Martín-Francés, L., Arsuaga, J. L., de Castro, J. M. B., Crété, L., ... & Stringer, C. (2022). Comparing the Boxgrove and Atapuerca (Sima de los Huesos) human fossils: Do they represent distinct paleodemes?. *Journal of Human Evolution*, *172*, 103253.

Lombao, D., Guardiola, M., & Mosquera, M. (2017). Teaching to make stone tools: New experimental evidence supporting a technological hypothesis for the origins of language. *Scientific Reports*, *7*(1), 14394.

Lombao, D., Rabuñal, J. R., Morales, J. I., Ollé, A., Carbonell, E., & Mosquera, M. (2022). The technological behaviours of Homo antecessor: core management and reduction intensity at Gran Dolina-TD6. 2 (Atapuerca, Spain). *Journal of Archaeological Method and Theory*, 1-38.

Lucas, A. J., Kings, M., Whittle, D., Davey, E., Happé, F., Caldwell, C. A., & Thornton, A. (2020). The value of teaching increases with tool complexity in cumulative cultural evolution. *Proceedings of the Royal Society B*, *287*(1939), 20201885.

Ludwig, B.V. (1999). *A Technological Reassessment of East African Plio-Pleistocene Lithic Artifact Assemblages*. New Brunswick: Rutgers University.

Luedtke, B. E. (1992). *An Archaeologist's Guide to Chert and Flint*. Cotsen Institute of Archaeology Press.

de Lumley, H., Barsky, D., & Cauche, D. (2009). Archaic stone industries from East Africa and southern Europe Pre-Oldowan and Oldowan. *The Cutting Edge. New Approaches to the Archaeology of Human Origins*. Stone Age Institute Publication Series.

Luncz, L. V., Mundry, R., & Boesch, C. (2012). Evidence for cultural differences between neighboring chimpanzee communities. *Current Biology*, *22*(10), 922-926.

Luncz, L. V., Arroyo, A., Falótico, T., Quinn, P., & Proffitt, T. (2022a). A primate model for the origin of flake technology. *Journal of Human Evolution*, *171*, 103250.

Luncz, L. V., Braun, D. R., Marreiros, J., Bamford, M., Zeng, C., Pacome, S. S., ... & Carvalho, S. (2022b). Chimpanzee wooden tool analysis advances the identification of percussive technology. *iScience*, 105315.

Lycett, S. J., & Eren, M. I. (2013). Levallois lessons: the challenge of integrating mathematical models, quantitative experiments and the archaeological record. *World Archaeology*, *45*(4), 519-538.

Lycett, S. J., & Gowlett, J. A. (2008). On questions surrounding the Acheulean 'tradition'. *World Archaeology*, *40*(3), 295-315.

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

M

Marshall-Pescini, S., & Whiten, A. (2008a). Chimpanzees (*Pan troglodytes*) and the question of cumulative culture: an experimental approach. *Animal Cognition*, *11*(3), 449–456.

Marshall-Pescini, S., & Whiten, A. (2008). Social learning of nut-cracking behavior in East African sanctuary-living chimpanzees (*Pan troglodytes schweinfurthii*). *Journal of Comparative Psychology*, *122*(2), 186.

Marzke, M. W., & Shackley, M. S. (1987). Hominid hand use in the Pliocene and Pleistocene: evidence from experimental archaeology and comparative morphology. *Journal of Human Evolution*, *15*(6), 439-460.

Marzke, M. W., Toth, N., Schick, K., Reece, S., Steinberg, B., Hunt, K., ... & An, K. N. (1998). EMG study of hand muscle recruitment during hard hammer percussion manufacture of Oldowan tools. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, *105*(3), 315-332.

Masi, S., Pouydebat, E., San-Galli, A., Meulman, E., Breuer, T., Reeves, J., & Tennie, C. (2022). Free hand hitting of stone-like objects in wild gorillas. *Scientific Reports*, *12*(1), 1-10.

Matsuzawa, T. (1994). Field experiments on use of stone tools by chimpanzees in the wild. *Chimpanzee Cultures*, 351-370.

McGrew, W. C. (1992). Tool-use by free-ranging chimpanzees: the extent of diversity. *Journal of Zoology, 228*(4), 689-694.

McGrew, W. C., Falótico, T., Gumert, M. D., & Ottoni, E. B. (2019). A simian view of the Oldowan: Reconstructing the evolutionary origins of human technology. In K. A. Overmann & F. L. Coolidge (Eds.), *Squeezing Minds from Stones: Cognitive Archaeology & the Evolution of the Human Mind* (pp. 13–41). Oxford University Press.

McNabb, J., Binyon, F., & Hazelwood, L. (2004). The large cutting tools from the South African Acheulean and the question of social traditions. *Current Anthropology*, *45*(5), 653-677.

McPherron, S. P., Alemseged, Z., Marean, C. W., Wynn, J. G., Reed, D., Geraads, D., ... & Béarat, H. A. (2010). Evidence for stone-tool-assisted consumption of animal tissues before 3.39 million years ago at Dikika, Ethiopia. *Nature*, *466*(7308), 857-860.

Mercader, J., Panger, M., & Boesch, C. (2002). Excavation of a chimpanzee stone tool site in the African rainforest. *Science*, *296*(5572), 1452-1455.

Mercader, J., Barton, H., Gillespie, J., Harris, J., Kuhn, S., Tyler, R., & Boesch, C. (2007). 4,300-year-old chimpanzee sites and the origins of percussive stone technology. *Proceedings of the National Academy of Sciences*, *104*(9), 3043-3048.

Mgeladze, A., Lordkipanidze, D., Moncel, M. H., Despriee, J., Chagelishvili, R., Nioradze, M., & Nioradze, G. (2011). Hominin occupations at the Dmanisi site, Georgia, Southern Caucasus: Raw materials and technical behaviours of Europe's first hominins. *Journal of Human Evolution*, *60*(5), 571-596.

Militky, J., & Kovacic, V. (1996). Ultimate mechanical properties of basalt filaments. *Textile Research Journal*, *66*(4), 225-229.

Mine, J. G., Slocombe, K. E., Willems, E. P., Gilby, I. C., Yu, M., Thompson, M. E., ... & Machanda, Z. P. (2022). Vocal signals facilitate cooperative hunting in wild chimpanzees. *Science Advances, 8*(30), eabo5553.

Mithen, S. (1996). Social learning and cultural tradition: interpreting Early Palaeolithic technology. *The Archaeology of Human Ancestry: Power, Sex and Tradition*, 207-229.

Montrey, M., & Shultz, T. R. (2020). The evolution of high-fidelity social learning. *Proceedings of the Royal Society B*, *287*(1928), 20200090.

Moore, M. W. (2020). Hominin stone flaking and the emergence of 'top-down' design in human evolution. *Cambridge Archaeological Journal*, *30*(4), 647-664.

Moore, M. W., & Brumm, A. (2009). *Homo floresiensis* and the African Oldowan. *Interdisciplinary Approaches to the Oldowan*, 61–69.

Moore, M. W., & Perston, Y. (2016). Experimental insights into the cognitive significance of early stone tools. *PLoS ONE, 11, e0158803*.

Moore, M. W., Sutikna, T., Morwood, M. J., & Brumm, A. (2009). Continuities in stone flaking technology at Liang Bua, Flores, Indonesia. *Journal of Human Evolution*, *57*(5), 503-526.

Mora, R., & De la Torre, I. (2005). Percussion tools in Olduvai Beds I and II (Tanzania): implications for early human activities. *Journal of Anthropological Archaeology*, *24*(2), 179-192.

Morgan, B., Eren, M. I., Khreisheh, N., Hill, G., Bradley, B., Jennings, T., & Smallwood, A. (2015). Clovis bipolar lithic reduction at Paleo Crossing, Ohio: A reinterpretation based on the examination of experimental replications. In J. Thomas & S. Ashley (Eds.), *Clovis: Current perspectives on technology, chronology, and adaptations* (pp. 121–143). College Station: Texas A&M Press.

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

Motes-Rodrigo, A., & Tennie, C. (2021). The Method of Local Restriction: In search of potential great ape culture-dependent forms. *Biological Reviews*, *96*(4), 1441–1461.

Motes-Rodrigo, A., Majlesi, P., Pickering, T. R., Laska, M., Axelsen, H., Minchin, T. C., ... & Hernandez-Aguilar, R. A. (2019). Chimpanzee extractive foraging with excavating tools: experimental modeling of the origins of human technology. *PLoS ONE*, *14*(5), e0215644.

Motes-Rodrigo, A., McPherron, S. P., Archer, W., Hernandez-Aguilar, R. A., & Tennie, C. (2022). Experimental investigation of orangutans' lithic percussive and sharp stone tool behaviours. *PLoS ONE*, *17*(2), e0263343.

Muller, A., & Clarkson, C. (2016). Identifying major transitions in the evolution of lithic cutting edge production rates. *PLoS ONE, 11*(12), e0167244.

Muller, A., Shipton, C., & Clarkson, C. (2022). Stone toolmaking difficulty and the evolution of hominin technological skills. *Scientific Reports*, *12*(1), 1-12.

Muto, G. R. (1974). A Proposed Model for Ideocultural Analysis of Chipped Stone Implements. 39th Annual Meeting of the Society for American Archaeology, Washington, D.C.

N

Neadle, D., Allritz, M., & Tennie, C. (2017). Food cleaning in gorillas: social learning is a possibility but not a necessity. *PLoS ONE*, *12*(12), e0188866.

Neadle, D., Bandini, E., & Tennie, C. (2020). Testing the individual and social learning abilities of task-naïve captive chimpanzees (Pan troglodytes sp.) in a nut-cracking task. *PeerJ*, *8*, e8734.

Neadle, D. L., Chappell, J., Clay, Z., & Tennie, C. (2021). Even under majority influence, great apes fail to copy novel actions. *OSF Preprints*. https://doi.org/10.31219/osf.io/swt9b

Neldner, K., Reindl, E., Tennie, C., Grant, J., Tomaselli, K., & Nielsen, M. (2020). A cross-cultural investigation of young children's spontaneous invention of tool use behaviours. *Royal Society Open Science*, *7*(5), 192240.

Neubauer, S., Hublin, J. J., & Gunz, P. (2018). The evolution of modern human brain shape. *Science Advances*, *4*(1), eaao5961.

Neufuss, J., Humle, T., Cremaschi, A., & Kivell, T. L. (2016). Nut-cracking behaviour in wild-born, rehabilitated bonobos (*Pan paniscus*): a comprehensive study of hand-preference, hand grips and efficiency. *American Journal of Primatology*, *79*(2), e22589.

Nielsen, M. (2012). Imitation, pretend play, and childhood: Essential elements in the evolution of human culture?. *Journal of Comparative Psychology*, *126*(2), 170.

Nishiaki, Y. (2019). Mastering hammer use in stone knapping: an experiment. In *Learning Among Neanderthals and Palaeolithic Modern Humans* (pp. 59-76). Springer, Singapore.

Nonaka, T., Bril, B., & 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*, 155–167.

Nowell, A., & White, M. (2010). Growing up in the Middle Pleistocene. *Stone Tools.* Boulder: University Press of Colorado, 67-82.

O

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

P

Pacheco, F. G., Guzmán, F. J. G., López, J. M. G., Pérez, A. S., Finlayson, C., Vidal, J. R., ... & Fa, D. A. (2012). The tools of the last Neanderthals: Morphotechnical characterisation of the lithic industry at level IV of Gorham's Cave, Gibraltar. *Quaternary international*, *247*, 151-161.

Page, S. N. (2014). Cultural transmission of lithic artefact traditions: an experimental approach (doctoral dissertation, University College London).

Panger, M. A., Brooks, A. S., Richmond, B. G., & Wood, B. (2002). Older than the Oldowan? Rethinking the emergence of hominin tool use. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, *11*(6), 235-245.

Pargeter, J., & Eren, M. I. (2017). Quantifying and Comparing Bipolar Versus Freehand Flake Morphologies, Production Currencies, and Reduction Energetics During Lithic Miniaturization. *Lithic Technology*, *42*(2–3), 90–108.

Pargeter, J., & Shea, J. J. (2019). Going big versus going small: Lithic miniaturization in hominin lithic technology. *Evolutionary Anthropology: Issues, News, and Reviews*, *28*(2), 72-85.

Pargeter, J., Khreisheh, N., & Stout, D. (2019). Understanding stone tool-making skill acquisition: Experimental methods and evolutionary implications. *Journal of Human Evolution*, *133*, 146–166.

Pargeter, J., Khreisheh, N., Shea, J. J., & Stout, D. (2020). Knowledge vs. know-how? Dissecting the foundations of stone knapping skill. *Journal of Human Evolution*, *145*, 102807.

Pargeter, J., Stout, D., Cheng, L., Kilgore, M. B. (2021) Testing the social, cognitive, and motor foundations of Paleolithic skill reproduction. *OSF*. doi:10.17605/OSF.IO/YEBRU

Parry, R., Dietrich, G., & Bril, B. (2014). Tool use ability depends on understanding of functional dynamics and not specific joint contribution profiles. *Frontiers in Psychology*, *5*.

Pelcin, A. W. (1997). 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*(12), 1107-1113.

Perreault, C. (2019). The quality of the archaeological record. In *The Quality of the Archaeological Record*. University of Chicago Press.

Piperno, M., Collina, C., Gallotti, R., Raynal, J. P., Kieffer, G., Bourdonnec, F. X. L., Poupeau, G., & Geraads, D. (2009). Obsidian exploitation and utilization during the Oldowan at Melka Kunture (Ethiopia). *Interdisciplinary Approaches to the Oldowan*, 111–128.

Placi, S. (2022). Sensitivity to geometry in humans and other animals. *In&Vertebrates*.

Ponce de León, M. S., Bienvenu, T., Marom, A., Engel, S., Tafforeau, P., Alatorre Warren, J. L., ... & Zollikofer, C. P. (2021). The primitive brain of early Homo. *Science*, *372*(6538), 165-171.

Pope, S. M., Taglialatela, J. P., Skiba, S. A., & Hopkins, W. D. (2018). Changes in frontoparietotemporal connectivity following do-as-I-do imitation training in chimpanzees (*Pan troglodytes*). *Journal of Cognitive Neuroscience*, *30*(3), 421-431.

Pradhan, G. R., Tennie, C., & van Schaik, C. P. (2012). Social organization and the evolution of cumulative technology in apes and hominins. *Journal of Human Evolution*, *63*(1), 180-190.

Proffitt, T., Luncz, L. V., Falótico, T., Ottoni, E. B., de la Torre, I., & Haslam, M. (2016). Wild monkeys flake stone tools. *Nature*, *539*(7627), 85-88.

Proffitt, T., Haslam, M., Mercader, J. F., Boesch, C., & Luncz, L. V. (2018). Revisiting Panda 100, the first archaeological chimpanzee nut-cracking site. *Journal of Human Evolution*, *124*, 117-139.

Proffitt, T., Reeves, J. S., Benito-Calvo, A., Sánchez-Romero, L., Arroyo, A., Malaijivitnond, S., & Luncz, L. V. (2021a). Three-dimensional surface morphometry differentiates behaviour on primate percussive stone tools. *Journal of the Royal Society Interface*, *18*(184), 20210576.

Proffitt, T., Bargalló, A., & de la Torre, I. (2021b). The effect of raw material on the identification of knapping skill: A case study from Olduvai Gorge, Tanzania. *Journal of Archaeological Method and Theory*, *29*(1), 50-82.

Proffitt, T., Reeves, J. S., Braun, D. R., Malaivijitnond, S., & Luncz, L. V. (2022a) Stone tool using macaques shed light on the emergence of hominin technology. *European Society for the Study of Human Evolution (ESHE) 2022*.

Proffitt, T., Reeves, J. S., Pacome, S. S., & Luncz, L. V. (2022b). Identifying functional and regional differences in chimpanzee stone tool technology. *Royal Society Open Science*, *9*(9), 220826.

Putt, S. S. (2015). The origins of stone tool reduction and the transition to knapping: An experimental approach. *Journal of Archaeological Science: Reports*, *2*, 51–60.

Putt, S. S., Woods, A. D., & Franciscus, R. G. (2014). The Role of Verbal Interaction During Experimental Bifacial Stone Tool Manufacture. *Lithic Technology*, *39*(2), 96–112.

Putt, S. S., Wijeakumar, S., Franciscus, R. G., & Spencer, J. P. (2017). The functional brain networks that underlie Early Stone Age tool manufacture. *Nature Human Behaviour*, *1*(6), 0102.

Putt, S. S., Wijeakumar, S., & Spencer, J. P. (2019). Prefrontal cortex activation supports the emergence of early stone age toolmaking skill. *NeuroImage*, *199*, 57-69.

Putt, S. S., Anwarzai, Z., Holden, C., Ruck, L., & Schoenemann, P. T. (2022). The evolution of combinatoriality and compositionality in hominid tool use: a comparative perspective. *International Journal of Primatology*, 1-46.

R

Ranhorn, K. L. (2017) Cultural transmission and lithic technology in Middle Stone Age eastern Africa (doctoral dissertation, The George Washington University).

Reader, S. M., & Laland, K. N. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proceedings of the National Academy of Sciences*, *99*(7), 4436-4441.

Reader, S. M., & Laland, K. N. (Eds.). (2003). *Animal innovation* (Vol. 10). Oxford: Oxford University Press.

Reeves, J. S., Proffitt, T., & Luncz, L. V. (2021). Modeling a primate technological niche. *Scientific Reports*, *11*(1), 1-9.

Rein, R., Bril, B., & Nonaka, T. (2013). Coordination strategies used in stone knapping: Coordination Strategies Used in Stone Knapping. *American Journal of Physical Anthropology*, *150*(4), 539–550.

Rein, R., Nonaka, T., & Bril, B. (2014). Movement Pattern Variability in Stone Knapping: Implications for the Development of Percussive Traditions. *PLoS ONE*, *9*(11), e113567.

Reindl, E., Beck, S. R., Apperly, I. A., & Tennie, C. (2016). Young children spontaneously invent wild great apes' tool-use behaviours. *Proceedings of the Royal Society B: Biological Sciences*, *283*, 20152402.

Reindl, E., Apperly, I. A., Beck, S. R., & Tennie, C. (2017). Young children copy cumulative technological design in the absence of action information. *Scientific Reports*, *7*(1), 1-11.

Reindl, E., Bandini, E., & Tennie, C. (2018). The zone of latent solutions and its relation to the classics: Vygotsky and Köhler. In *Evolution of Primate Social Cognition* (pp. 231-248). Springer, Cham.

Reindl, E., Tennie, C., Apperly, I. A., Lugosi, Z., & Beck, S. R. (2022). Young children spontaneously invent three different types of associative tool use behaviour. *Evolutionary Human Sciences*, *4*.

Renfrew, C. (2004). Towards a theory of material engagement. In E. DeMarrais, C. Gosden, & C. Renfrew (Eds.) *Rethinking Materiality: The Engagement of Mind with the Material World* (pp. 23-32). Cambridge: McDonald Institute for Archaeological Research.

Rezek, Z., Lin, S. C., & Dibble, H. L. (2016). The role of controlled experiments in understanding variation in flake production. In *Archaeological Variability and Interpretation in Global Perspective* (pp. 307-320). University press of Colorado.

Režek, Ž., Dibble, H. L., McPherron, S. P., Braun, D. R., & Lin, S. C. (2018*).* Two million years of flaking stone and the evolutionary efficiency of stone tool technology. *Nature Ecology & Evolution, 2*, 628-633.

Robbins, M. M., Ando, C., Fawcett, K. A., Grueter, C. C., Hedwig, D., Iwata, Y., ... & Yamagiwa, J. (2016). Behavioral variation in gorillas: evidence of potential cultural traits. *PLoS ONE*, *11*(9), e0160483.

Rolian, C., & Carvalho, S. (2017). Tool use and manufacture in the last common ancestor of *Pan* and *Homo*. In M. N. Muller, R. W. Wrangham, & D. R. Pilbeam (Eds.), *Chimpanzees and Human Evolution* (pp. 602–644). The Belknap Press of Harvard University Press.

Rolian, C., Lieberman, D. E., & Zermeno, J. P. (2011). Hand biomechanics during simulated stone tool use. *Journal of Human Evolution*, *61*(1), 26–41.

Roux, V., Bril, B., & Dietrich, G. (1995). Skills and learning difficulties involved in stone knapping: The case of stone-bead knapping in Khambhat, India. *World Archaeology*, *27*(1), 63-87.

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Rutz, C., Klump, B. C., Komarczyk, L., Leighton, R., Kramer, J., Wischnewski, S., ... & Masuda, B. M. (2016). Discovery of species-wide tool use in the Hawaiian crow. *Nature*, *537*(7620), 403-407.

S

Sahnouni, M., & van der Made, J. (2009). The Oldowan in North Africa within a biochronological framework. *The Cutting Edge: New Approaches to the Archaeology of Human Origins*, *3*.

Saldana, C., Fagot, J., Kirby, S., Smith, K., & Claidière, N. (2019). High-fidelity copying is not necessarily the key to cumulative cultural evolution: a study in monkeys and children. *Proceedings of the Royal Society B*, *286*(1904), 20190729.

Samuni, L., Wegdell, F., & Surbeck, M. (2020). Behavioural diversity of bonobo prey preference as a potential cultural trait. *eLife*, *9*, e59191.

Sasaki, T., & Biro, D. (2017). Cumulative culture can emerge from collective intelligence in animal groups. *Nature Communications*, *8*(1), 1-6.

Sayers, K., & Lovejoy, C. O. (2008). The chimpanzee has no clothes: a critical examination of *Pan troglodytes* in models of human evolution. *Current Anthropology*, *49*(1), 87-114.

Schick, K. D. (1994). The Movius Line reconsidered: perspectives on the earlier Paleolithic of eastern Asia. In R. S. Corruccini, & R. L. Cochon (Eds.), *Integrative Paths to the Past* (pp. 569-96). Englewood Cliffs: Prentice Hall.

Schick, K. D., & Toth, N. (1994). *Making Silent Stones Speak: Human Evolution and the Dawn of Technology*. Simon & Schuster.

Schick, K. D., Toth, N., Garufi, G., Savage-Rumbaugh, E. S., Rumbaugh, D., & Sevcik, R. (1999). Continuing investigations into the stone tool-making and tool-using capabilities of a bonobo (*Pan paniscus*). *Journal of Archaeological Science*, *26*(7), 821–832.

Schillinger, K., Mesoudi, A., & Lycett, S. J. (2014). Copying error and the cultural evolution of “additive” vs. “reductive” material traditions: an experimental assessment. *American Antiquity*, *79*(1), 128-143.

Schillinger, K., Mesoudi, A., & Lycett, S. J. (2015). The impact of imitative versus emulative learning mechanisms on artifactual variation: Implications for the evolution of material culture. *Evolution and Human Behavior*, *36*(6), 446–455.

Schöning, C., Humle, T., Möbius, Y., & McGrew, W. C. (2008). The nature of culture: technological variation in chimpanzee predation on army ants revisited. *Journal of Human Evolution*, *55*(1), 48-59.

Schuppli, C., & van Schaik, C. P. (2019). Animal cultures: how we've only seen the tip of the iceberg. *Evolutionary Human Sciences*, *1*.

Schürch, B., Wettengl, S., Fröhle, S., Conard, N., & Schmidt, P. (2022). The origin of chert in the Aurignacian of Vogelherd Cave investigated by infrared spectroscopy. *PLoS ONE*, *17*(8), e0272988.

Semaw, S. (2006). The Oldest Stone Artifacts From Gona (2.6-2.5 Ma), Afar, Ethiopia: Implications for Understanding the Earliest Stages of Stone Knapping. In K. Schick, & N. Toth (Eds.), *The Oldowan: Case Studies Into the Earliest Stone Age* (pp. 43-75). Gosport: Stone Age Institute Press.

Semaw, S., Rogers, M. J., Quade, J., Renne, P. R., Butler, R. F., Dominguez-Rodrigo, M., ... & Simpson, S. W. (2003). 2.6-Million-year-old stone tools and associated bones from OGS-6 and OGS-7, Gona, Afar, Ethiopia. *Journal of Human Evolution*, *45*(2), 169-177.

Semaw, S., Rogers, M. J., Simpson, S. W., Levin, N. E., Quade, J., Dunbar, N., ... & Everett, M. (2020). Co-occurrence of Acheulian and Oldowan artifacts with *Homo erectus* cranial fossils from Gona, Afar, Ethiopia. *Science Advances*, *6*(10), eaaw4694.

Shea, J. J. (2006). Child's play: reflections on the invisibility of children in the Paleolithic record. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, *15*(6), 212-216.

Shea, J. J. (2010). Stone Age visiting cards revisited: A strategic perspective on the lithic technology of early hominin dispersal. In *Out of Africa I* (pp. 47-64). Springer, Dordrecht.

Shea, J. J. (2013). Lithic modes A–I: a new framework for describing global-scale variation in stone tool technology illustrated with evidence from the East Mediterranean Levant. *Journal of Archaeological Method and Theory*, *20*(1), 151-186.

Shea, J. J. (2015). Making and using stone tools: advice for learners and teachers and insights for archaeologists. *Lithic Technology*, *40*(3), 231-248.

Shea, J. J. (2017). Occasional, obligatory, and habitual stone tool use in hominin evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, *26*(5), 200-217.

Shea, J. J. (2019). European Upper Palaeolithic cultural taxa: better off without them?. *Antiquity*, *93*(371), 1359-1361.

Shea, J. J. (2022). Chasing Mirages: Seeking Standardization among Prehistoric Stone Tools. *Lithic Technology*, 1-8.

Sheets, P. D., & Muto, G. R. (1972). Pressure blades and total cutting edge: an experiment in lithic technology. *Science*, *175*(4022), 632-634.

Shelley, P. H. (1990). Variation in lithic assemblages: an experiment. *Journal of Field Archaeology*, *17*(2), 187-193.

Shilton, D. (2019). Is Language Necessary for the Social Transmission of Lithic Technology?. *Journal of Language Evolution*, *4*(2), 124-133.

Shipton, C. (2010). Imitation and shared intentionality in the Acheulean. *Cambridge Archaeological Journal*, *20*(2), 197-210.

Shipton, C. (2020). The unity of Acheulean culture. In *Culture History and Convergent Evolution* (pp. 13-27). Springer, Cham.

Shipton, C., & Nielsen, M. (2015). Before Cumulative Culture: The Evolutionary Origins of Overimitation and Shared Intentionality. *Human Nature*, *26*(3), 331–345.

Speer, C. (2018). Using Porcelain Replicas for Precision Control in Flintknapping Experiments. *Advances in Archaeological Practice, 6*(1), 72-81.

Snyder, W. D., Reeves, J. S., & Tennie, C. (2022). Early knapping techniques do not necessitate cultural transmission. *Science Advances*, *8*(27), eabo2894.

Stade, C. M. (2017). *Lithic Morphological Variability as a Proxy for Palaeolithic Linguistic Ability: A Knapping Training Study Exploring Cultural Transmission, Theory of Mind, and Language* (doctoral dissertation, University of Southampton).

Stahl, J. (2008). Who Were the Flintknappers? A Study of Individual Characteristics. *Lithic Technology, 33*(2), 161–172.

Sterelny, K. (2020). Innovation, life history and social networks in human evolution. *Philosophical Transactions of the Royal Society B*, *375*(1803), 20190497.

Sterelny, K. (2021). Foragers and their tools: risk, technology and complexity. *Topics in Cognitive Science*, *13*(4), 728-749.

Sterelny, K., & Hiscock, P. (in press). Cumulative Culture, Archaeology, and the Zone of Latent Solutions. *Current Anthropology*.

Sternke, F., & Sørensen, M. (2007). The identification of children's flint knapping products in Mesolithic Scandinavia. In S. McCartan, R. Schulting, G. Warren, P. C. Woodman, (Eds.) *Mesolithic Horizons. Papers Presented at the Seventh International Conference on the Mesolithic in Europe, Belfast 2005*. Oxbow Books.

Stout, D. (2002). Skill and cognition in stone tool production: An ethnographic case study from Irian Jaya. *Current Anthropology*, *43*(5), 693–722.

Stout, D. (2005). Neural foundations of perception and action in stone knapping. In V. Roux, & B. Bril (Eds.) *Stone Knapping: The Necessary Conditions for a Uniquely Hominin Behaviour*. Oxford: Oxbow books.

Stout, D. (2006). Oldowan toolmaking and hominin brain evolution: theory and research using Positron Emission Tomography (PET). In N. Toth & K. Schick (Eds.) *The Oldowan: Case Studies into the Earliest Stone Age* (pp. 267-306). Gosport: Stone Age Institute Press.

Stout, D., & Semaw, S. (2006). Knapping skill of the earliest stone toolmakers: Insights from the study of modern human novices. In N. Toth & K. Schick (Eds.), *The Oldowan: Case Studies into the Earliest Stone Age* (pp. 307–320). Gosport: Stone Age Institute Press.

Stout, D., & Chaminade, T. (2007). The evolutionary neuroscience of tool making. *Neuropsychologia*, *45*(5), 1091–1100.

Stout, D., & Chaminade, T. (2009). Making Tools and Making Sense: Complex, Intentional Behaviour in Human Evolution. *Cambridge Archaeological Journal*, *19*(1), 85–96.

Stout, D., & Chaminade, T. (2012). Stone tools, language and the brain in human evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1585), 75–87.

Stout, D., & Hecht, E. (2015). Neuroarchaeology. *Human Paleoneurology*, 145–175.

Stout, D., & Hecht, E. E. (2017). Evolutionary neuroscience of cumulative culture. *Proceedings of the National Academy of Sciences*, *114*(30),

Stout, D., & Khreisheh, N. (2015). Skill Learning and Human Brain Evolution: An Experimental Approach. *Cambridge Archaeological Journal*, *25*(4), 867–875.

Stout, D., Quade, J., Semaw, S., Rogers, M. J., & Levin, N. E. (2005). Raw material selectivity of the earliest stone toolmakers at Gona, Afar, Ethiopia. *Journal of Human Evolution*, *48*(4), 365–380.

Stout, D., Toth, N., Schick, K., & Chaminade, T. (2008). Neural correlates of Early Stone Age toolmaking: Technology, language and cognition in human evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1499), 1939–1949.

Stout, D., Schick, K., & Toth, N. (2009). Understanding Oldowan knapping skill: An experimental study of skill acquisition in modern humans. In N. Toth & K. Schick (Eds.), *The Cutting Edge: New Approaches in the Archaeology of Human Origins* (pp. 247–266). Stone Age Institute Press.

Stout, D., Semaw, S., Rogers, M. J., & Cauche, D. (2010). Technological variation in the earliest Oldowan from Gona, Afar, Ethiopia. *Journal of Human Evolution*, *58*(6), 474–491.

Stout, D., Passingham, R., Frith, C., Apel, J., & Chaminade, T. (2011). Technology, expertise and social cognition in human evolution: Technology and cognition in human evolution. *European Journal of Neuroscience*, *33*(7), 1328–1338.

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

Stout, D., Hecht, E., Khreisheh, N., Bradley, B., & Chaminade, T. (2015). Cognitive demands of Lower Paleolithic toolmaking. *PLOS ONE*, *10*(4), e0121804.

Stout, D., Rogers, M. J., Jaeggi, A. V., & Semaw, S. (2019). Archaeology and the origins of human cumulative culture: A case study from the earliest Oldowan at Gona, Ethiopia. *Current Anthropology*, *60*(3), 309–340.

Stout, D., Chaminade, T., Apel, J., Shafti, A., & Faisal, A. A. (2021). The measurement, evolution, and neural representation of action grammars of human behavior. *Scientific Reports*, *11*(1), 13720.

Sykes, R. W. (2021). *Kindred: Neanderthal Life, Love, Death and Art*. Van Haren Publishing.

T

Tebbich, S., Taborsky, M., Fessl, B., & Blomqvist, D. (2001). Do woodpecker finches acquire tool-use by social learning?. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*(1482), 2189-2193.

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

Tennie, C., & Call, J. (in press). Unmotivated subjects cannot provide interpretable data and tasks with sensitive learning periods require appropriately aged subjects. A Commentary on Koops et al.'s "Field experiments find no evidence that chimpanzee nut cracking can be independently motivated". *Animal Behavior and Cognition.*

Tennie, C., Hedwig, D., Call, J., & Tomasello, M. (2008). An experimental study of nettle feeding in captive gorillas. *American Journal of Primatology: Official Journal of the American Society of Primatologists*, *70*(6), 584-593.

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: On the evolution of cumulative culture. *Phil. Trans. R. Soc. B*, *364*, 2405–2415.

Tennie, C., Call, J., & Tomasello, M. (2010). Evidence for emulation in chimpanzees in social settings using the floating peanut task. *PLoS One*, *5*(5), e10544.

Tennie, C., Call, J., & Tomasello, M. (2012). Untrained chimpanzees (*Pan troglodytes schweinfurthii*) fail to imitate novel actions. *PLoS ONE, 7*(8), e41548.

Tennie, C., Braun, D. R., Premo, L. S., & McPherron, S. P. (2016). The Island Test for Cumulative Culture in the Paleolithic. In M. N. Haidle, N. J. Conard, & M. Bolus (Eds.), *The Nature of Culture* (pp. 121–133). Springer Netherlands.

Tennie, C., Premo, L. S., Braun, D. R., & McPherron, S. P. (2017). Early Stone Tools and Cultural Transmission: Resetting the Null Hypothesis. *Current Anthropology*, *58*(5), 652–672.

Tennie, C., Bandini, E., van Schaik, C. P., & Hopper, L. M. (2020a). The zone of latent solutions and its relevance to understanding ape cultures. *Biology & Philosophy*, *35*(5), 55.

Tennie, C., Hopper, L. M., & van Schaik, C. P. (2020b). On the origin of cumulative culture: Consideration of the role of copying in culture-dependent traits and a reappraisal of the Zone of Latent Solutions Hypothesis. In *Chimpanzees in Context* (pp. 428-453). University of Chicago Press.

Thornton, A., & Raihani, N. J. (2010). Identifying teaching in wild animals. *Learning & Behavior*, *38*(3), 297-309.

Thornton, A., Samson, J., & Clutton-Brock, T. (2010). Multi-generational persistence of traditions in neighbouring meerkat groups. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1700), 3623–3629. https://doi.org/10.1098/rspb.2010.0611

Tomasello, M. (1999). The human adaptation for culture. *Annual Review of Anthropology*, *28*(1), 509–529.

Tomasello, M. (2003). *The Cultural Origins of Human Cognition*. Harvard Univ. Press.

Tomasello, M., & Call, J. (2004). The role of humans in the cognitive development of apes revisited. *Animal Cognition*, *7*(4), 213-215.

Tomasello, M., Savage-Rumbaugh, S., & Kruger, A. C. (1993). Imitative learning of actions on objects by children, chimpanzees, and enculturated chimpanzees. *Child Development*, *64*(6), 1688-1705.

Tomasello, M., Call, J., & Gluckman, A. (1997). Comprehension of novel communicative signs by apes and human children. *Child Development*, 1067-1080.

Torres, C., & Preysler, J. B. (2020). Experts Also Fail: a New Methodological Approach to Skills Analysis in Lithic Industries. *Journal of Paleolithic Archaeology*, *3*(4), 889–917.

Toth, N. P. (1982). *The Stone Technologies of Early Hominids at Koobi Fora, Kenya: An Experimental Approach*. University of California, Berkeley.

Toth, N. (1985). The Oldowan reassessed: A close look at early stone artifacts. *Journal of Archaeological Science*, *12*(2), 101–120.

Toth, N., & Schick, K. (1994). Early stone industries and inferences regarding language and cognition. In K.R. Gibson, & T. Ingold (Eds.), *Tools, Language and Cognition in Human Evolution* (pp. 346-362). Cambridge University Press.

Toth, N., & Schick, K. (2009). The Oldowan: the tool making of early hominins and chimpanzees compared. *Annual Review of Anthropology*, *38*, 289-305.

Toth, N., & Schick, K. (2011). Factors affecting variability in Early Stone Age lithic assemblages: Personal observations from actualistic studies. *Casting the Net Wide: Papers in Honor of Glynn Isaac and His Approach to Human Origins Research*, 53–74.

Toth, N., & Schick, K. (2018). An overview of the cognitive implications of the Oldowan Industrial Complex. *Azania: Archaeological Research in Africa*, *53*(1), 3–39.

Toth, N., Schick, K. D., Savage-Rumbaugh, E. S., Sevcik, R. A., & Rumbaugh, D. 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*(1), 81–91.

Toth, N., Schick, K., & Semaw, S. (2006). A comparative study of the stone tool-making skills of *Pan*, *Australopithecus*, and *Homo sapiens*. In K. Schick, & N. Toth (Eds.), *The Oldowan: Case Studies Into the Earliest Stone Age* (pp. 155-222). Gosport: Stone Age Institute Press.

Tramacere, A., & Moore, R. (2018). Reconsidering the role of manual imitation in language evolution. *Topoi*, *37*(2), 319-328.

Tylor, E. B., & Suith, E. L. D. E. R. (1871). Primitive culture: Researches into the development of mythology, Philosophy. *The Athenaeum*, 261.

V

Vale, G. L., Davis, S. J., Lambeth, S. P., Schapiro, S. J., & Whiten, A. (2017). Acquisition of a socially learned tool use sequence in chimpanzees: implications for cumulative culture. *Evolution and Human Behavior*, *38*(5), 635-644.

van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C. D., Singleton, I., ... & Merrill, M. (2003). Orangutan cultures and the evolution of material culture. *Science*, *299*(5603), 102-105.

van Schaik, C. P., Pradhan, G. R., & Tennie, C. (2019). Teaching and curiosity: sequential drivers of cumulative cultural evolution in the hominin lineage. *Behavioral Ecology and Sociobiology*, *73*(1), 1-11.

Visalberghi, E. (1987). Acquisition of nut-cracking behaviour by 2 capuchin monkeys (*Cebus apella*). *Folia Primatologica*, *49*(3-4), 168-181.

W

Watson, P. J. (1995). Archaeology, anthropology, and the culture concept. *American Anthropologist*, *97*(4), 683-694.

Westergaard, G. C., & Suomi, S. J. (1993). Use of a tool-set by capuchin monkeys (Cebus apella). *Primates*, *34*(4), 459-462.

Westergaard, G. C., & Suomi, S. J. (1994a). A simple stone-tool technology in monkeys. *Journal of Human Evolution*, *27*(5), 399-404.

Westergaard, G. C., & Suomi, S. J. (1994b). The use and modification of bone tools by capuchin monkeys. *Current Anthropology*, *35*(1), 75-77.

Westergaard, G. C., & Suomi, S. J. (1995a). The stone tools of capuchins (*Cebus apella*). *International Journal of Primatology*, *16*(6), 1017-1024.

Westergaard, G. C., & Suomi, S. J. (1995b). The manufacture and use of bamboo tools by monkeys: possible implications for the development of material culture among East Asian hominids. *Journal of Archaeological Science*, *22*(5), 677-681.

Whiten, A. (2015). Experimental studies illuminate the cultural transmission of percussive technologies in *Homo* and *Pan*. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1682), 20140359.

Whiten, A. (2016). The evolution of hominin culture and its ancient pre-hominin foundations. In *The Nature of Culture* (pp. 27-39). Springer, Dordrecht.

Whiten, A. (2017). Culture extends the scope of evolutionary biology in the great apes. *Proceedings of the National Academy of Sciences*, *114*(30), 7790-7797.

Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., ... & Boesch, C. (1999). Cultures in chimpanzees. *Nature*, *399*(6737), 682-685.

Whiten, A., Horner, V., & Marshall-Pescini, S. (2003). Cultural panthropology. *Evolutionary Anthropology: Issues, News, and Reviews*, *12*(2), 92–105.

Whiten, A., Horner, V., Litchfield, C. A., & Marshall-Pescini, S. (2004). How do apes ape?. *Animal Learning & Behavior*, *32*(1), 36-52.

Whiten, A., McGuigan, N., Marshall-Pescini, S., & Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2417-2428.

Whiten, A., Schick, K., & Toth, N. (2009b). The evolution and cultural transmission of percussive technology: Integrating evidence from palaeoanthropology and primatology. *Journal of Human Evolution*, *57*(4), 420–435.

Whittaker, J. C. (1987). Making arrowpoints in a prehistoric Pueblo. *Lithic Technology*, *16*(1), 1-12.

Whittaker, J. C. (1994). *Flintknapping: Making and Understanding Stone Tools*. University of Texas Press.

Williams, E. M., Gordon, A. D., & Richmond, B. G. (2010). Upper limb kinematics and the role of the wrist during stone tool production. *American Journal of Physical Anthropology*, *143*(1), 134–145.

Williams-Hatala, E. M., Hatala, K. G., Gordon, M., Key, A., Kasper, M., & Kivell, T. L. (2018). The manual pressures of stone tool behaviors and their implications for the evolution of the human hand. *Journal of Human Evolution*, *119*, 14–26.

Williams-Hatala, E. M., Hatala, K. G., Key, A., Dunmore, C. J., Kasper, M., Gordon, M., & Kivell, T. L. (2020). Kinetics of stone tool production among novice and expert tool makers. *American Journal of Physical Anthropology*.

Wright, R. V. (1972). Imitative learning of a flaked stone technology-the case of an orangutan. *The Australian Journal of Anthropology*, *8*(4), 296.

Wynn, T. (1979). The Intelligence of Later Acheulean Hominids. *Man*, *14*(3), 371.

Wynn, T. (1985). Piaget, stone tools and the evolution of human intelligence. *World Archaeology*, *17*(1), 32-43.

Wynn, T. (1988). Tools and the evolution of human intelligence. In R. W. Byrne & A. Whiten (Eds.), *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans* (pp. 271–284). Clarendon Press/Oxford University Press.

Wynn, T. (1991). Tools, Grammar and the Archaeology of Cognition. *Cambridge Archaeological Journal*, *1*(2), 191–206.

Wynn, T., & Coolidge, F. L. (2004). The expert Neandertal mind. *Journal of Human Evolution*, *46*(4), 467-487.

Wynn, T., & Coolidge, F. L. (2016). Archeological insights into hominin cognitive evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, *25*(4), 200-213.

Wynn, T., & McGrew, W. C. (1989). An ape's view of the Oldowan. *Man*, *24*(3), 383.

Wynn, T., Hernandez-Aguilar, R. A., Marchant, L. F., & McGrew, W. C. (2011). “An ape's view of the Oldowan” revisited. *Evolutionary Anthropology: Issues, News, and Reviews*, *20*(5), 181-197.

Wynn, T., Berlant, T., Overmann, K. A., & Coolidge, F. L. (2019). The handaxe aesthetic. *Squeezing Minds from Stones: Cognitive Archaeology and the Evolution of the Human Mind*, 278-303.

Y

Yaroshevsky, A. A. (2006). Abundances of chemical elements in the Earth’s crust. *Geochemistry International*, *44*(1), 48-55.

Z

Ponce De León, M. S., & Zollikofer, C. P. (2001). The brain and its case: computer-based case studies on the relation between software and hardware in living and fossil hominid skulls. *The Brain and its Case: Computer-based Case Studies on the Relation Between Software and Hardware in Living and Fossil Hominid Skulls*, 1000-1006.

Zorrilla-Revilla, G., Vidal-Cordasco, M., Prado-Nóvoa, O., & Terradillos-Bernal, M. (2021). Know-how, or how knapping experience can affect a prehistoric lithic workshop. *Lithic Technology*, *46*(3), 221-235.

Zuberbühler, K., Gygax, L., Harley, N., & Kummer, H. (1996). Stimulus enhancement and spread of a spontaneous tool use in a colony of long-tailed macaques. *Primates*, *37*(1), 1-12.

Appendix I: Supplementary methods

Explanation of candidate contribution to collaborative work

This chapter is based on collaborative work. The text contains (modified) excerpts from the main and supplementary text of the following publication:

Snyder, W.D., Reeves, J.S., & Tennie, C. (2022). Early knapping techniques do not necessitate cultural transmission. *Science Advances, 8*(27), eabo2894. doi:10.1126/sciadv.abo2894

Conceptualization: C.T., W.D.S., J.S.R. Investigation: W.D.S. Visualization: W.D.S. Formal analysis: W.D.S., C.T. Writing (original draft): W.D.S. Writing (review and editing): W.D.S., C.T.

Some of the remaining content is from as-yet unpublished, in-preparation work:

Snyder, W.D., Reeves, J.S., & Tennie, C. (in prep). Untitled naïve human toolmaking and tool use paper.

Participation

Inclusion in the study was conditional upon the provision of informed consent, whereby participants were given sufficient time to read all information sheets before signing the relevant consent forms.

Since we could not know *a priori* how many technique-naive (and totally naïve) participants we would encounter during data collection, we tested a relatively large sample of participants (thirty participants, before dropouts). Combined with the four-hour testing sessions, we hypothesized that this would prove sufficient to result in the logically necessary amount of data – i.e., that at least one participant would re-innovate at least one knapping technique during testing, while being naïve at time of innovation to this know-how (based on prior power analysis in Bandini et al., 2021c for a similar study on chimpanzees). Conducting these experiments is taxing and resource-intensive on several levels, meaning that testing further participants was considered logistically unsound (this would have been impossible anyway due to the repeated onset of further COVID-19-related restrictions).

Testing Apparatus and Procedure

Video recording of the testing sessions was performed using one Sony Handycam HDR-CX250 (Camera 2), one Sony Handycam HDR-CX330 (Camera 3), and one Sony Handycam HDR-CX450 (Camera 1), all mounted on tripods situated on spots pre-designated (with duct tape) in the testing space. Recording using the camcorders was done so at the widest possible angle setting for the lenses, and the centrally located camcorder (Camera 2) was enhanced with a wide-angle lens attachment to capture as much action as possible in the testing space. The whole experiment was recorded entirely without audio due to data protection constraints set by ZENDAS, the Central Data Protection Authority for Universities in Baden-Württemberg. The microphone jacks for the camcorders were physically obstructed to ensure that no audio was recorded (this was once tested and verified to work before testing began).

The introduction to the apparatus included some clarification of the mechanics of the tendon box (i.e., that its rope prevents its door from being opened, but without using the words "rope" or "door"). The experimenter used gesturing and general language towards the glass blank, river cobble, and granite block – rather than specifically naming the objects. This means the participants were not verbally cued to the properties of the material or any specific potential uses for the objects. Experimenters therefore never provided any instruction about stone toolmaking or use, nor were participants allowed to elicit any help in this regard from the experimenters. Experimenters were required to follow a script to ensure the same vague terminology and phrasing was used for all participant sessions.

Upon each successful solving of the problem, the participant gained access inside the puzzle box to a reward-token - a paper slip representing the monetary reward (printed on one side). New reward tokens were placed inside the box after each success, and the box then re-set. The first reward value was always 10 €, but values varied thereafter. Collected reward values across all successes were added together at the end of the session and the sum then paid out to the participant in legal tender.

The determination of the reward value in each session and with each attempt followed a formula of diminishing returns, since the initial motivation was assumed to be high and, in an ecological setting, the first “portion(s)” of a raw material or food source would likely be the most valuable and additional retrievals (and the effort involved) would be decreasingly profitable over time. The curve to the diminishing returns formula that was used has a long tail (with low value rewards) as the overall success rate of participants could not be predicted and the total money available for each session (per the research grant ERC Starting Grant STONECULT) was limited. Participants were guaranteed to earn 10 € for their first success. Afterwards, the new reward would be selected from the deck of tokens and then placed face-down in the rope apparatus (only revealed to the participant after they succeed). Each time a token was drawn from the deck of notes, a note with 0,01€ was added into the deck and the full deck was shuffled (thus reducing the expected reward value with each iteration; participants were not made aware this was occurring). The reward “deck” contained the followed distribution of rewards:

- 10 € x 1
- 2 € x 10
- 1 € x 20
- 0,50 € x 40
- 0,10 € x 50
- 0,01 € x ∞ (theoretically only)

Exhaustion of cores applied in cases where participants stopped engaging with the test material (e.g., actively percussing or pounding the glass or stone, pursuing repeated or new solutions for the tendon box apparatus) for more than five minutes at a time. If and when a participant exhausted a core, they were provided a new one by the experimenters. Participants were allowed a maximum of one new blank for each hour of the experiment (e.g., if a participant received their second core at one hour and twenty minutes into the session, they were allowed to receive a new core - at the earliest - two hours and twenty minutes into the study).

If participants deemed their final core exhausted before the end of the last hour of the study, they had the option of waiting idly (receiving full reimbursement for the hour) or of leaving (receiving only money for the time they waited and forgoing any additional reward, even if their core was not truly exhausted). This particular situation did not come to fruition in the study.

Data

The behaviors were separated into bouts, which were either single actions or sequences of actions with a recorded start and end point. The start of a bout was defined by the engagement of the participant with the raw materials and puzzle box in any way that might lead to toolmaking or opening of the box. The end of a bout was defined when the character of behavior changed (thus starting a separate, new bout, e.g., the participant fluidly switched from passive hammer technique to bipolar technique) or the participant ceased engaging in the particular behavior for thirty seconds or longer.

Coding of toolmaking techniques was determined through an “elements-based” approach (adapted from previous approaches) (Key & Lycett, 2014), which was determined as a reliable way to define the behaviors actually performed by the study participants (since many of these behaviors did not readily fit in pre-existing classificatory models of EST production techniques) while also being agnostic to both the mechanic and outcome of any fracture events. This approach was also conceived as applicable to further scenarios involving humans and non-human animals, making cross-context and cross-species comparisons more viable. In this outcome-neutral approach, there are four classes of elements that can be present in the constellation of objects and actions that take place in toolmaking: active, passive, auxiliary, and target elements. Active elements are objects in motion (i.e., imbued with/possessing kinetic energy), and passive elements are the primary recipient of force from an active element. Auxiliary elements are primarily used for stabilization or support to a passive element but can also be transformed by the toolmaking process or even be mechanically active in a secondary sense to the active element. Target elements are, in essence, the target of a tool use behavior (e.g., the rope of the tendon box or in ecological scenarios of extractive foraging: nuts or animal carcasses).

The identification of these elements was used to categorize early knapping techniques (de lat Torre, 2019; Harmand et al., 2015; Putt, 2015; Toth & Schick, 2011, 2018; Wynn & McGrew, 1989) as: freehand knapping, passive hammer technique, bipolar technique, and projectile technique (with additional coding for simultaneous toolmaking and use).

A hypothesis-unaware person (J.K.) was hired as a student research assistant with the assignment of performing secondary coding of a subset of the data (all behavioral clips from 7 participants, 25% of participant sample). The reliability coder used the same coding protocol as the experimenter W.D.S. Furthermore, this individual was untrained in the methods and theory of lithic technology and thus considered as unbiased towards traditional methods of understanding toolmaking and the techniques thereof.

Participants were ranked for their naivety on the basis of their responses to the Questionnaire on the Prior Experience of Study Participants. We used the following ranks (categories) to characterize the amount and nature of experience individuals had with stone tools:

0. Participant is totally naive

1. Participant has conceptual knowledge of stone tools (i.e., conscious of their existence)

2. Participant has seen a stone tool or a depiction of a stone tool

3. Participant has seen the production of a stone tool [here, technique know-how could have been imparted on participants]

4. Participant has direct hands-on experience with knapping and, with it, at least one knapping technique

Labels contained the code name of the participant, the number of the blank (labelled sequentially from 1 to n), and whether it was successful or unsuccessful to access a reward.

Detached pieces were identified as either flakes or angular fragments. We followed the same definition used by Braun et al. (2019), whereby flakes show all the hallmarks of conchoidal fracture (i.e., platform and bulb of percussion) while angular fragments were those artefacts that lacked hallmarks of conchoidal fracture.

The validity of our claims regarding knapping depends upon the accurate identification of the artefacts as flakes (i.e., produced by conchoidal fracture in the same manner as the prehistoric equivalents). Typing of detached pieces was performed by W.D.S. and then repeated by a hypothesis-unaware second observer (A.F.). A.F. is an experienced lithic analyst, but – as alluded to – was not informed of the purpose of the typing and the theoretical and methodological framework of the experiments. The second observer was tasked with identifying a randomly selected subsample of 155 detached pieces (approximately 10% of the 1553 detached pieces from all participants) as either flakes or angular fragments.

For enhanced clarity about the categories, illustrations of toolmaking behaviors were created in Adobe Illustrator by tracing over screen captures of video recordings.

Lithic illustrations are a standard practice in archaeological research and analysis (Cerasoni, 2021). Cores were covered in non-reflective spray (due to transparency and reflectiveness of glass) and scanned using an ARTEC Spider while flakes were photographed with a digital camera, as attempts at 3D scanning produced an unclosed mesh with noise along the fine, cutting edge of the glass flakes. 3D scans of cores were imported into Blender, converted into orthographic projection (this mapped the technological information of the 3D cores onto 2D space without distortion) and then screen-capped (using an internally developed protocol serving as a “digital photography studio”). Attributes of the artefacts were traced in Adobe Illustrator to generate lithic illustrations according to the STIVA method (Cerasoni, 2019).

Appendix II: Supplementary results

Explanation of candidate contribution to collaborative work

This chapter is based on collaborative work. The text contains (modified) excerpts from the main and supplementary text of the following publication:

Snyder, W.D., Reeves, J.S., & Tennie, C. (2022). Early knapping techniques do not necessitate cultural transmission. *Science Advances, 8*(27), eabo2894. doi:10.1126/sciadv.abo2894

Conceptualization: C.T., W.D.S., J.S.R. Investigation: W.D.S. Visualization: W.D.S. Formal analysis: W.D.S., C.T. Writing (original draft): W.D.S. Writing (review and editing): W.D.S., C.T.

Some of the remaining content is from as-yet unpublished, in-preparation work:

Snyder, W.D., Reeves, J.S., & Tennie, C. (in prep). Untitled naïve human toolmaking and tool use paper.

Behaviors and toolmaking techniques

Confirmed toolmaking events were more stereotyped and less variable in terms of their technique categorization, with five individuals using one technique exclusively. 12 individuals used two of the different technique categories to successfully fracture objects, while 9 individuals used three of the different technique categories, and one individual implemented all four technique categories for toolmaking. However, the difference in distributions of technique preferences between potential and confirmed toolmaking was not statistically significant (Pearson's chi-squared test, X^2=6.3808, p=0.1725).

The first potential toolmaking event was most often categorized as freehand technique (n=10, 35.7%) followed by both passive hammer and bipolar (for each: n=7, 25.0%) and then projectile (n=4, 14.3%). The distribution shifted when considering only events with successful fractures, with bipolar technique as the most common (n=10, 37.0%), followed by passive hammer and freehand (for each, n=8, 29.6%) and projectile again the least common (n=1, 3.7%). The difference in distributions between potential and confirmed toolmaking, however, was not statistically significant (Pearson's chi-squared test, X^2=2.601, p=0.4573).

Interrater reliability using Cohen's kappa tests determined that video coding of toolmaking and techniques in the selected subsample of trials (n = 228) was reliable between the two observers using both counts of relevant bouts in each trial ($\kappa > 0.6$) and presence/absence of relevant bouts in each trial ($\kappa > 0.9$).

Material outcomes

Only 2 individuals did not produce any objects that could be classified as flakes (P2 and P23), though in both these cases, the glass hemispheres showed signs of removals by conchoidal fracture. Most individuals produced assemblages that could be characterized as flake-dominated (i.e., >50% of artefacts were conchoidal flakes). There were a total of 1172 flakes (73.3%), 361 angular fragments (22.6%), 33 glass cores (2.1%), and 33 river cobbles (as cores, percussors, or hammer-cores; 2.1%).

The hypothesis-unaware second observer (A.F.) typed 122 of 155 (78.71%) of the detached pieces as flakes, compared with 126 flakes out of 155 artefacts (81.29%) for the original observer, W.D.S. The second observer identified 7 objects as flakes that the original observer did not, two of which were non-glass objects. The original observer identified 11 objects as flakes (all made of glass) that the second observer identified as angular fragments. The Cohen's kappa test determined that the flake identification protocol was reliable in the selected subsample ($\kappa = 0.637$).

There were no statistical differences in terms of EPA between any naivety groups and between all naivety groups and the Oldowan reference material.

There were no statistical differences in terms of PD between the naivety groups from the experiment. However, each pairwise comparison for PD between experimental naivety groups and the Oldowan reference material was significant.

All pairwise comparisons used a Wilcoxon rank sum test, and no *p*-value adjustment was applied.

Table S1. Demographic and experiential information.

Individual	Sex	Decade of Birth	Naivety Level
P1	F	1990s	3
P2	F	1990s	2
P3	M	1990s	2
P4	F	1990s	2
P5	M	2000s	2
P6	F	1990s	2
P7	M	2000s	1
P8	M	1990s	2
P9	M	1960s	2
P10	F	1970s	2
P11	F	1990s	0
P12	F	2000s	2
P13	M	1970s	2
P14	M	1980s	0
P15	M	1990s	2
P16	M	1970s	2
P17	F	1960s	2
P19	F	2000s	4
P21	M	2000s	2
P22	M	1990s	2
P23	F	2000s	2
P24	M	1980s	2
P25	F	1980s	2
P26	F	1990s	2
P27	M	1990s	2
P28	F	1960s	3
P29	F	1960s	2
P30	M	2000s	2

Table S2. Innovation of toolmaking behavior.

Individual	Naivety	Did participant make a cutting tool? (y/n)	Did participant make *and* use a cutting tool? (y/n)	First potential toolmaking event				First confirmed toolmaking event				*n* of successful solutions (i.e. opening tendon box)
				Mins. elapsed	*n* of prior solution attempts	Trial #	TM Teq.	Mins. elapsed	*n* of prior solution attempts	Trial #	TM Teq.	
P1	3	Y	Y	1	0	1	PH	1	0	1	PH	35
P2	2	Y	Y	5	7	3	FH	29	8	4	PH	4
P3	2	Y	Y	2	0	1	BP	2	0	1	BP	54
P4	2	Y	Y	14	2	3	FH	37	12	5	PH	51
P5	2	Y	Y	3	2	1	PJ	3	2	1	PJ	73
P6	2	Y	Y	98	56	2	FH	238	75	6	PH	6
P7	1	Y	Y	7	4	2	PJ	25	16	4	BP	55
P8	2	Y	Y	1	1	1	PH	1	1	1	PH	70
P9	2	Y	Y	9	1	2	FH	9	1	2	FH	60
P10	2	Y	Y	26	15	5	FH	31	16	6	BP	38
P11	0	Y	Y	31	9	4	BP	31	9	4	BP	57
P12	2	Y	Y	1	1	1	FH	10	7	2	FH	72
P13	2	Y	Y	17	15	3	PH	20	15	3	FH	52
P14	0	Y	Y	21	10	1	FH	27	11	2	FH	57
P15	2	Y	Y	17	15	3	BP	17	16	3	BP	54
P16	2	Y	Y	70	29	6	FH	77	31	8	FH	45
P17	2	Y	Y	4	6	1	PH	4	6	1	PH	55
P19	4	Y	Y	10	5	2	PH	26	28	2	BP	55
P21	2	Y	N	79	69	4	PJ	194	115	7	BP	7
P22	2	Y	Y	8	2	2	BP	8	2	2	BP	27
P23	2	Y	N	17	17	1	BP	n.a.	n.a.	n.a.	n.a.	7
P24	2	Y	Y	32	19	3	FH	32	19	3	FH	70
P25	2	Y	Y	45	3	1	PH	47	31	3	PH	57
P26	2	Y	Y	6	5	2	BP	6	5	2	BP	59
P27	2	Y	Y	2	1	1	FH	3	1	1	FH	53
P28	3	Y	Y	1	0	1	PH	1	0	1	PH	64
P29	2	Y	Y	41	31	4	PJ	82	53	6	FH	59
P30	2	Y	Y	10	2	3	BP	10	2	3	BP	78
Mean				20.6 ±25.2	11.7 ±16.7	2.3 ±1.4	n.a.	36.0 ±56.3	17.9 ±26.1	3.1 ±2.0	n.a.	49.1 ±21.1

Table S3. Technique of first potential and confirmed toolmaking events.

Event	Passive Hammer		Bipolar		Freehand		Projectile	
	n	*%*	*n*	*%*	*n*	*%*	*n*	*%*
First potential toolmaking event	7	25.0%	7	25.0%	10	35.7%	4	14.3%
First confirmed toolmaking event	8	29.6%	10	37.0%	8	29.6%	1	3.7%

Table S4. Preferences for techniques in potential and confirmed toolmaking.

Condition	Passive Hammer		Bipolar		Freehand		Projectile		Anvil-oriented		Hand-oriented		Opportunistic	
	n	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
Potential TM events	7	25.0	10	35.7	4	14.3	0	0	5	17.9	0	0	2	7.1
TM events with observable fracture	10	37.0	13	48.1	3	11.1	0	0	0	0	0	0	1	3.7

Table S5. Distribution of techniques for potential toolmaking events.

Individual	*N* of TM	Passive Hammer		Bipolar		Freehand		Projectile		Preference
		n	%	*n*	%	*n*	%	*n*	%	
P1	41	41	92.7	3	7.3	0	0	0	0	PH
P2	9	4	44.4	1	11.1	1	11.1	3	33.3	AO
P3	50	3	6.0	47	94.0	0	0	0	0	BP
P4	59	36	61.0	14	23.7	8	13.6	1	1.7	PH
P5	41	40	97.6	0	0	0	0	1	2.4	PH
P6	25	5	20.0	3	12.0	17	68.0	0	0	FH
P7	57	24	42.1	4	7.0	28	49.1	1	1.8	OP
P8	44	1	2.3	2	4.5	41	93.2	0	0	FH
P9	62	4	6.5	32	51.6	26	41.9	0	0	BP
P10	34	3	8.8	25	73.5	5	14.7	1	2.9	BP
P11	38	1	2.6	36	94.7	1	2.6	0	0	BP
P12	34	0	0	32	94.1	2	5.9	0	0	BP
P13	74	16	21.6	54	73.0	4	5.4	0	0	BP
P14	58	54	93.1	1	1.7	3	5.2	0	0	PH
P15	87	26	29.9	42	48.3	7	8.0	12	13.8	AO
P16	47	21	44.7	20	42.6	6	12.8	0	0	AO
P17	73	57	78.1	9	12.3	7	9.6	0	0	PH
P19	56	2	3.6	54	96.4	0	0	0	0	BP
P21	25	2	8.0	12	48.0	0	0	11	44.0	AO
P22	153	72	47.1	55	35.9	26	17.0	0	0	AO
P23	16	0	0	16	100.0	0	0	0	0	BP
P24	46	0	0	23	50.0	23	50.0	0	0	OP
P25	52	52	100.0	0	0	0	0	0	0	PH
P26	77	67	87.0	3	3.9	4	5.2	3	3.9	PH
P27	65	11	16.9	33	50.8	21	32.3	0	0	BP
P28	68	20	29.4	48	70.6	0	0	0	0	BP
P29	47	0	0	0	0	44	93.6	3	6.4	FH
P30	142	0	0	5	3.5	136	95.8	1	0.7	FH
Overall	1580	559	35.4	574	36.3	410	25.9	37	2.3	
Mean	56.429	20.071	n.a.	20.5	n.a.	14.643	n.a.	1.321	n.a.	
Std. dev.	31.635	23.090	n.a.	19.346	n.a.	27.063	n.a.	3.031	n.a.	

Table S6. Frequency of coded potential toolmaking (TM) events. % is the percentage of total bouts represented by a category. κ1 is the Cohen's kappa coefficient calculated using exact counts of the relevant bouts per trial. κ2 is the Cohen's kappa coefficient calculated using Boolean data (presence/absence of the specific bout per trial). Where fracture and severing of the puzzle box rope occurred simultaneously, the event was coded as toolmaking and use (TMU).

	n	%	$\bar{x}$	x_{min}	x_{max}	κ_1	κ_2
TM bouts	1580	100.0	56.4	9 (P2)	153 (P22)	0.707	0.978
TMU bouts	260	16.5	9.3	0	79 (P30)	0.538	0.721
Freehand	410	25.9	14.6	0	136 (P30)	0.632	0.939
Passive hammer	559	35.4	20.1	0	72 (P22)	0.756	0.991
Bipolar	574	36.3	20.5	0	55 (P22)	0.932	0.966
Projectile	37	2.3	1.3	0	12 (P15)	1.000	1.000

Table S7. Distribution of techniques for confirmed toolmaking events.

Individual	*N* of TM	Passive Hammer		Bipolar		Freehand		Projectile		Preference
		n	%	*n*	%	*n*	%	*n*	%	
P1	39	37	94.9	2	5.1	0	0	0	0	PH
P2	5	4	80.0	1	20.0	0	0	0	0	PH
P3	49	3	6.1	46	93.9	0	0	0	0	BP
P4	35	26	74.3	9	25.7	0	0	0	0	PH
P5	40	39	97.5	0	0	0	0	1	2.5	PH
P6	2	2	100.0	0	0	0	0	0	0	PH
P7	36	17	47.2	2	5.6	17	47.2	0	0	OP
P8	41	1	2.4	1	2.4	39	95.1	0	0	FH
P9	49	2	4.1	32	65.3	15	30.6	0	0	BP
P10	23	0	0	22	95.7	1	4.3	0	0	BP
P11	36	0	0	36	100.0	0	0	0	0	BP
P12	33	0	0	32	97.0	1	3.0	0	0	BP
P13	56	6	10.7	49	87.5	1	1.8	0	0	BP
P14	54	53	98.1	0	0	1	1.9	0	0	PH
P15	58	17	29.3	31	53.4	6	10.3	4	6.9	BP
P16	28	7	25.0	17	60.7	4	14.3	0	0	BP
P17	57	47	82.5	8	14.0	2	3.5	0	0	PH
P19	50	0	0	50	100.0	0	0	0	0	BP
P21	6	1	16.7	5	83.3	0	0	0	0	BP
P22	49	28	57.1	19	38.8	2	4.1	0	0	PH
P24	31	0	0	22	71.0	9	29.0	0	0	BP
P25	46	46	100.0	0	0	0	0	0	0	PH
P26	45	39	86.7	3	6.7	3	6.7	0	0	PH
P27	46	5	10.9	29	63.0	12	26.1	0	0	BP
P28	66	18	27.3	48	72.7	0	0	0	0	BP
P29	43	0	0	0	0	43	100.0	0	0	FH
P30	72	0	0	2	2.8	70	97.2	0	0	FH
Overall	1095	398	36.3	466	42.6	226	20.6	5	0.5	
Mean	40.6	19.9	n.a.	21.2	n.a.	14.1	n.a.	2.5	n.a.	
Std. dev.	17.2	18.0	n.a.	17.5	n.a.	19.8	n.a.	2.1	n.a.	

Table S8. Artefact type distributions for the assemblages.

		Total	Flakes		Angular Fragments		Glass Cores		River Cobbles		Percussive Objects
			n	*%*	*n*	*%*	*n*	*%*	*n*	*%*	
Individual	P1	39	26	66.7	11	28.2	1	2.6	1	2.6	1
	P2	3	0	0.0	1	33.3	1	33.3	1	33.3	1
	P3	114	76	66.7	36	31.6	1	0.9	1	0.9	1
	P4	43	35	81.4	6	14.0	1	2.3	1	2.3	1
	P5	38	23	60.5	13	34.2	1	2.6	1	2.6	1
	P6	5	2	40.0	1	20.0	1	20.0	1	20.0	2
	P7	41	26	63.4	13	31.7	1	2.4	1	2.4	2
	P8	76	45	59.2	29	38.2	1	1.3	1	1.3	1
	P9	64	46	71.9	16	25.0	1	1.6	1	1.6	1
	P10	45	35	77.8	8	17.8	1	2.2	1	2.2	2
	P11	41	35	85.4	4	9.8	1	2.4	1	2.4	1
	P12	67	53	79.1	11	16.4	2	3.0	1	1.5	1
	P13	104	85	81.7	17	16.3	1	1.0	1	1.0	1
	P14	77	61	79.2	14	18.2	1	1.3	1	1.3	2
	P15	75	42	56.0	30	40.0	1	1.3	2	2.7	2
	P16	83	43	51.8	37	44.6	1	1.2	2	2.4	2
	P17	124	105	84.7	15	12.1	3	2.4	1	0.8	1
	P19	95	77	81.1	14	14.7	3	3.2	1	1.1	1
	P21	8	2	25.0	4	50.0	1	12.5	1	12.5	1
	P22	80	63	78.8	14	17.5	1	1.3	2	2.5	3
	P23	4	0	0.0	2	50.0	1	25.0	1	250%	1
	P24	41	37	90.2	2	4.9	1	2.4	1	2.4	1
	P25	60	48	80.0	10	16.7	1	1.7	1	1.7	2
	P26	30	26	86.7	2	6.7	1	3.3	1	3.3	2
	P27	45	39	86.7	4	8.9	1	2.2	1	2.2	1
	P28	108	60	55.6	44	40.7	1	0.9	3	2.8	2
	P29	57	53	93.0	2	3.5	1	1.8	1	1.8	1
	P30	32	29	90.6	1	3.1	1	3.1	1	3.1	1
	Mean	57.1±33.6	41.9±25.7		12.9±12.0		1.2±0.5		1.2±0.5		
Naivety Level	0	118	96	81.4	18	15.3	2	1.7	2	1.7	3
	1	41	26	63.4	13	31.7	1	2.4	1	2.4	2
	2	1198	887	74.0	261	21.8	25	2.1	25	2.1	30
	3	147	86	58.5	55	37.4	2	1.4	4	2.7	3
	4	95	77	81.1	14	14.7	3	3.2	1	1.1	1
	Total	1599	1172	73.3	361	22.6	33	2.1	33	2.1	39

Table S9. Means and standard deviations for the measurements of flakes (including mass, maximum dimension, width, thickness, platform depth and exterior platform angle). Values are reported according to individual, full sample, and naivety level.

		n	Mass (g)		Max. Dim. (mm)		Width (mm)		Thickness (mm)		PD (mm)		EPA (°)	
			$\bar{x}$	*s*	$\bar{x}$	*s*	$\bar{x}$	*s*	$\bar{x}$	*s*	$\bar{x}$	*s*	$\bar{x}$	*s*
Individual	Z1	26	1.8	2.3	30.9	16.1	16.0	7.7	2.7	1.2	2.6	n.a.	90.0	n.a.
	Z3	76	5.2	8.5	31.3	10.7	19.3	9.1	5.2	4.0	4.6	3.7	80.0	6.9
	Z4	35	2.5	4.0	26.6	12.9	17.1	8.7	3.2	2.0	3.3	1.7	77.2	4.8
	Z5	23	5.6	10.0	33.9	16.6	19.0	10.7	4.4	3.7	4.7	4.5	78.7	9.6
	Z6	2	3.9	4.9	30.1	12.7	23.8	18.1	4.4	1.6	1.1	n.a.	77.0	n.a.
	Z7	26	3.5	8.2	27.6	12.6	19.2	10.6	3.2	2.9	7.6	5.4	81.7	7.1
	Z8	45	7.8	10.9	35.3	13.4	21.3	11.6	6.0	4.4	5.0	4.5	73.1	5.8
	Z9	46	5.2	8.1	30.2	11.0	17.8	9.7	5.6	4.1	5.0	1.7	74.8	5.2
	Z10	35	7.8	13.1	37.4	15.1	22.5	11.9	5.2	3.9	5.1	4.3	70.4	10.5
	Z11	35	2.2	4.2	22.9	11.9	14.7	8.6	3.1	2.1	2.5	2.1	74.3	3.4
	Z12	53	11.3	21.1	34.0	16.8	20.8	13.0	7.6	7.3	8.0	8.0	71.6	11.1
	Z13	85	4.1	6.1	30.4	10.0	18.7	8.8	4.6	2.7	3.3	2.7	76.3	11.1
	Z14	61	1.9	3.5	25.8	10.5	14.9	7.7	3.5	2.0	1.8	n.a.	81.0	n.a.
	Z15	42	9.8	19.7	29.8	17.7	16.6	10.9	6.7	6.5	7.2	5.5	72.9	6.9
	Z16	43	8.1	15.1	32.5	14.4	19.0	11.2	6.4	5.4	5.4	3.0	66.2	7.2
	Z17	105	8.4	17.4	37.1	14.1	20.3	9.6	6.2	5.0	14.7	17.3	78.8	14.3
	Z19	77	6.1	12.0	31.5	10.9	20.4	10.5	5.1	4.9	4.2	4.7	76.2	7.1
	Z21	2	0.7	0.2	19.3	1.0	15.9	0.8	1.5	0.3	n.a.	n.a.	n.a.	n.a.
	Z22	63	4.7	7.7	30.6	11.9	16.7	9.4	5.2	3.8	6.3	4.6	76.3	5.0
	Z24	37	5.8	15.0	32.0	13.2	17.9	10.5	4.1	4.0	5.5	4.3	79.7	3.9
	Z25	48	2.1	3.0	26.7	11.0	15.3	8.1	4.0	1.9	4.8	2.4	82.3	11.5
	Z26	26	2.2	3.4	28.9	14.3	16.1	7.7	2.9	1.9	2.1	n.a.	78.0	n.a.
	Z27	39	1.0	1.6	22.3	8.5	12.1	6.5	2.7	1.3	2.1	0.2	77.5	6.4
	Z28	60	6.2	9.3	31.1	12.9	18.2	10.9	5.5	4.0	5.6	3.2	74.1	9.6
	Z29	53	4.1	5.9	32.0	9.1	20.1	8.7	4.4	2.5	2.0	1.1	78.4	8.0
	Z30	29	1.5	1.4	28.5	8.2	14.8	4.8	3.1	1.2	1.4	0.3	88.0	0.0
	Mean	1172	5.3	11.2	30.8	13.0	18.2	9.9	4.9	4.2	5.0	4.7	75.6	8.6
Naivety Level	0&1	122	2.3	5.0	25.4	11.4	15.7	8.7	3.3	2.2	3.7	3.7	76.7	5.4
	2	887	5.7	12.0	31.5	13.2	18.4	9.9	5.1	4.3	5.1	4.9	75.5	8.9
	3	86	4.9	8.1	31.0	13.9	18.0	10.2	4.7	3.6	5.5	3.2	74.8	9.9
	4	77	6.1	12.0	31.5	10.9	20.4	10.5	5.1	4.9	4.2	4.7	76.2	7.1

Table S10. Means and standard deviations for the measurements of angular fragments (including mass, maximum dimension, width, and thickness). Values are reported according to individual, full sample, and naivety level.

		n	Mass (g)		Max. Dim. (mm)		Width (mm)		Thickness (mm)	
			$\bar{x}$	*s*	$\bar{x}$	*s*	$\bar{x}$	*s*	$\bar{x}$	*s*
Individual	Z1	11	1.1	1.0	20.9	7.2	11.0	5.4	3.1	1.4
	Z2	1	0.1	n.a.	6.6	n.a.	4.4	n.a.	3.2	n.a.
	Z3	36	7.6	14.4	30.7	13.8	17.2	7.4	7.0	4.4
	Z4	6	0.4	0.2	21.5	8.0	9.6	2.9	2.3	0.8
	Z5	13	21.7	40.2	33.6	12.7	19.7	14.0	10.7	14.2
	Z6	1	16.8	n.a.	73.1	n.a.	17.9	n.a.	7.0	n.a.
	Z7	13	29.8	73.9	31.8	26.9	21.0	19.5	7.6	11.4
	Z8	29	2.0	2.0	25.4	6.2	14.1	6.8	4.8	2.7
	Z9	16	6.5	15.7	29.4	14.3	13.9	8.0	6.1	4.4
	Z10	8	0.4	0.2	21.8	3.6	10.1	4.3	1.8	1.4
	Z11	4	0.5	0.3	20.5	9.3	16.2	11.0	2.4	1.0
	Z12	11	3.0	4.7	27.0	7.6	13.3	8.3	5.1	3.8
	Z13	17	28.8	53.1	34.5	20.0	23.4	15.8	10.5	9.7
	Z14	14	1.0	0.6	20.2	5.1	10.3	3.1	3.8	1.3
	Z15	30	11.0	30.7	26.9	18.9	14.4	13.6	5.9	5.6
	Z16	37	13.9	22.7	35.4	18.9	19.6	11.3	8.6	6.1
	Z17	15	1.4	1.1	31.6	10.1	11.8	2.9	3.0	1.6
	Z19	14	1.8	1.4	24.0	5.4	13.6	3.8	6.0	3.7
	Z21	4	35.6	28.1	54.9	21.7	19.6	11.1	15.3	9.2
	Z22	14	17.1	55.4	31.5	20.9	18.4	12.7	6.4	8.2
	Z23	2	8.9	12.5	42.2	43.5	13.4	7.6	5.0	6.4
	Z24	2	1.9	0.9	28.6	5.8	13.3	2.3	6.2	4.5
	Z25	10	1.8	2.0	25.9	10.6	13.3	6.6	4.2	2.8
	Z26	2	0.3	0.1	20.5	4.3	7.9	1.4	1.3	0.0
	Z27	4	0.5	0.4	18.1	4.0	8.3	2.5	2.6	2.0
	Z28	44	7.0	15.1	25.6	11.1	15.6	9.0	7.9	6.0
	Z29	2	3.0	3.0	30.4	3.7	17.1	11.5	3.6	1.3
	Z30	1	0.5	n.a.	22.3	n.a.	10.2	n.a.	2.7	n.a.
	Mean	361	9.2	27.2	2.9	1.5	1.6	1.0	0.6	0.6
Naivety Level	0&1	31	13.0	49.0	25.1	18.5	15.5	13.9	5.2	7.6
	2	261	9.8	26.5	30.1	15.8	16.0	10.4	6.4	6.4
	3	55	5.8	13.7	24.6	10.5	14.7	8.5	6.9	5.7
	4	14	1.8	1.4	24.0	5.4	13.6	3.8	6.0	3.7

Table S11. The measurements of the impossible flakes used in the study (including mass, maximum dimension, width, and thickness). Artefact number (i.e. order in the session that artefacts were collected) and success number (i.e. when the impossible flake was used in the sequence of successful apparatus solution events) are also reported in this table.

	Artefact Nr.	Success Nr.	Mass (g)	Max. Dim. (mm)	Width (mm)	Thickness (mm)
Z6	1	3	16.8	73.1	17.9	7
Z21	1	5	25.6	73	15.8	11.6
Z23	1	4	17.7	72.9	18.7	9.5
Mean			20.1±4.8	73.0±0.1	17.5±1.5	9.4±2.3

Table S12. Distribution of angular fragments according to material categories (cobble, granite, and glass), referring to the original angular fragment was removed (or in the case of impossible flakes, simply glass).

		Total	Cobble		Granite		Glass	
			n	%	*n*	%	*n*	%
Individual	Z1	11	8	72.7%	0	0.0%	3	27.3%
	Z2	1	0	0.0%	1	100.0%	0	0.0%
	Z3	36	18	50.0%	0	0.0%	18	50.0%
	Z4	6	0	0.0%	1	16.7%	5	83.3%
	Z5	13	4	30.8%	0	0.0%	9	69.2%
	Z6	1	0	0.0%	0	0.0%	1	100.0%
	Z7	13	3	23.1%	0	0.0%	10	76.9%
	Z8	29	0	0.0%	0	0.0%	29	100.0%
	Z9	16	0	0.0%	0	0.0%	16	100.0%
	Z10	8	0	0.0%	0	0.0%	8	100.0%
	Z11	4	0	0.0%	0	0.0%	4	100.0%
	Z12	11	0	0.0%	0	0.0%	11	100.0%
	Z13	17	7	41.2%	0	0.0%	10	58.8%
	Z14	14	2	14.3%	5	35.7%	7	50.0%
	Z15	30	14	46.7%	6	20.0%	10	33.3%
	Z16	37	24	64.9%	0	0.0%	13	35.1%
	Z17	15	0	0.0%	0	0.0%	15	100.0%
	Z19	14	0	0.0%	0	0.0%	14	100.0%
	Z21	4	3	75.0%	0	0.0%	1	25.0%
	Z22	14	2	14.3%	0	0.0%	12	85.7%
	Z23	2	0	0.0%	0	0.0%	2	100.0%
	Z24	2	0	0.0%	0	0.0%	2	100.0%
	Z25	10	0	0.0%	0	0.0%	10	100.0%
	Z26	2	0	0.0%	0	0.0%	2	100.0%
	Z27	4	0	0.0%	0	0.0%	4	100.0%
	Z28	44	23	52.3%	2	4.5%	19	43.2%
	Z29	2	0	0.0%	0	0.0%	2	100.0%
	Z30	1	0	0.0%	0	0.0%	1	100.0%
	Mean	12.9±12.0						
Total		361	108	29.9%	15	4.2%	238	65.9%

Table S13. Attributes and metrics for glass cores reduced during the study.

Individual	Naivety	Core #	Artefact #	Mass (g)	Max. Dim. (mm)	Width (mm)	Thickness (mm)	Scar Count	Step Terminations	Step Terms. / Scar Count
Z1	3	C6	35	470	97.5	96.2	44.8	24	13	0.54
Z2	2	C7	2	576	98.6	98.6	46	n.a.	n.a.	
Z3	2	C1	49	465	95.6	79.3	45.9	15	10	0.67
Z4	2	C9	30	479	97.9	76.9	46.5	21	10	0.48
Z5	2	C10	9	392	95.5	80.8	46	20	11	0.55
Z6	2	C11	4	539	96.8	95.9	44	4	0	0.00
Z7	1	C12	30	442	96.7	78	45.8	16	10	0.63
Z8	2	C13	37	137	66.6	64.9	41.6	13	5	0.38
Z9	2	C14	32	129	74.1	67.1	30	18	9	0.50
Z10	2	C15	19	254	80.4	69.9	44.7	16	5	0.31
Z11	0	C16	36	449	95.1	85.2	44.8	17	9	0.53
Z12	2	C17	26	50	53.3	44.1	20.3	10	8	0.80
Z12	2	C4	35	302	97.8	78.6	46	7	3	0.43
Z13	2	C18	54	103	54.4	46.9	38.7	21	14	0.67
Z14	0	C19	55	350	89.9	86.6	44	21	14	0.67
Z15	2	C20	50	84	58.6	51.4	34.3	7	3	0.43
Z16	2	C21	24	129	59.5	58.3	36.1	13	7	0.54
Z17	2	C22	14	52	48	37.3	25.7	15	3	0.20
Z17	2	C5	30	60	55	43	28.9	11	3	0.27
Z17	2	C23	36	471	98.9	84.6	45.7	18	10	0.56
Z19	4	C25	36	28	42.5	34.5	20	15	4	0.27
Z19	4	C25	42	68	51.3	35.3	31.7	15	2	0.13
Z19	4	C26	54	480	98.4	78.6	47	15	1	0.07
Z21	2	C28	3	579	99	98.4	44.1	8	3	0.38
Z22	2	C29	29	119	60.2	41	36.3	20	5	0.25
Z23	2	C30	3	594	99.1	98.7	45.3	5	0	0.00
Z24	2	C31	22	325	86.7	62.6	43.9	18	11	0.61
Z25	2	C32	42	398	89.7	86.9	46.1	31	23	0.74
Z26	2	C33	15	481	96.7	90.8	45.2	27	13	0.48
Z27	2	C34	34	506	95.5	92.4	45.6	30	11	0.37
Z28	3	C35	68	91	62	48.8	31.8	16	6	0.38
Z29	2	C36	1	284	82.9	68.2	46.6	32	21	0.66
Z30	2	C37	1	437	90.9	88.9	43.9	29	16	0.55
mean				312.81±190.24	80.76±19.29	71.17±20.81	40.22±7.97	17.13±7.37	8.22±5.71	0.44±0.21

Table S14. Means and standard deviations for core metrics and attributes according to naivety levels.

Naivety	*n*	Mass (g)		Max. Dim. (mm)		Width (mm)		Thickness (mm)		Scar Count		Step Terms.		Scar Count / Step Terms.	
		$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s
0&1	3	413.7	55.7	93.9	3.6	83.3	4.6	44.9	0.9	18.0	2.6	11.0	2.6	0.6	0.1
2	25	317.8	191.3	81.3	18.3	72.2	19.9	40.7	7.5	17.6	8.1	8.5	6.0	0.5	0.2
3	2	280.4	268.3	79.8	25.1	72.5	33.5	38.3	9.2	20.0	5.7	9.5	4.9	0.5	0.1
4	3	192.2	249.9	64.1	30.1	49.5	25.2	32.9	13.5	15.0	0.0	2.3	1.5	0.2	0.1

Table S15. Attributes and metrics for river cobbles (as hammers or hammer-cores) used during the study.

Individual	Naivety	Cobble ID	Artefact #	Mass (g)	Max. Dim. (mm)	Width (mm)	Thickness (mm)	Scar Count
Z1	3	M23		339.77	78	72	49.7	0
Z2	2	S22		259.56	86.8	55.3	36.8	0
Z3	2	S49		339.5	86.9	64.8	43.8	0
Z4	2	S4		341.98	78.9	66.7	52.1	0
Z5	2	M34		460.76	102	60.3	45.2	0
Z6	2	M29		609.66	97.3	81.9	63.8	0
Z7	0&1	M27		486.46	114.1	77.5	40.1	0
Z8	2	M33		465.65	81	68.7	62.1	0
Z9	2	S28		314.05	77.3	58.8	53	0
Z10	2	C15		336.51	79.2	68.5	52.9	0
Z11	0&1	C16		526.77	85.9	73	60	0
Z12	2	S15		263.84	81.1	47.1	45.1	0
Z13	2	S50		266.92	725	555	511	0
Z14	0&1	S57		300.43	77.3	74.5	41.5	5
Z15	2	M35		273.93	81.7	50	49.1	3
Z15	2	M28	52	237.94	86.1	68.7	25.4	2
Z16	2	M20		381.85	84.4	63.3	49.3	3
Z16	2	M5		62.36	53.7	33.2	28.3	7
Z17	2	S27		290.71	79.2	54.1	45.4	0
Z19	4	M21		527.64	89.2	80.2	65.4	0
Z21	2	M12		364.48	72.7	60.5	58.7	1
Z22	2	C29		363.11	94.3	59.2	44.3	1
Z22	2	C29	31.2	253.3	96.2	55	43.6	2
Z23	2	S3	4	305.5	99.3	52.4	43	0
Z24	2	M18		634	106.5	69.1	60.7	0
Z25	2	M17		379.57	87.6	66.8	48.2	0
Z26	2	S38	17	259.82	68.3	61.4	42.5	0
Z27	2	M3		361.38	89.6	67	49.4	0
Z28	3	S34		266.62	68.2	57.9	52.4	0
Z28	3	M24	19	55.97	54.4	38.8	29.9	0
Z28	3	M24	8	158.32	66.8	46.5	41	5
Z29	2	M36		369.83	78.5	60.7	56.9	0
Z30	2	M15		377.33	98.7	51.9	44.5	0
mean				340.47±129.50	103.22±112.41	76.39±86.64	61.67±81.24	0.88±1.78

Table S16. Means and standard deviations for cobble metrics and attributes according to naivety levels.

Naivety	*n*	Mass (g)		Max. Dim. (mm)		Width (mm)		Thickness (mm)		Scar Count	
		$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s
0&1	3	437.9	120.7	92.4	19.3	75.0	2.3	47.2	11.1	1.7	2.9
2	25	342.9	116.2	110.9	128.5	80.0	99.4	66.2	93.1	0.8	1.6
3	4	205.2	124.3	66.9	9.7	53.8	14.4	43.3	10.1	1.3	2.5
4	1	527.6	n.a.	89.2	n.a.	80.2	n.a.	65.4	n.a.	n.a.	n.a.

Table S17. p-Values from pairwise comparisons of PD using Wilcoxon rank sum tests. Cells for statistically significant comparisons marked in grey.

Naivety Level	2	3&4	Oldowan (*46*)
0&1	0.14559	0.28869	0.00011
2		0.61145	$< 2 \times 10^{-16}$
3&4			2.1×10^{-9}

Table S18. p-Values from pairwise comparisons of EPA using Wilcoxon rank sum tests.

Naivety Level	2	3&4	Oldowan (*46*)
0&1	0.72	0.63	0.83
2		0.87	0.59
3&4			0.64

Table S19. p-Values from pairwise comparisons of total cutting successes per second using Wilcoxon rank-sum tests

Naivety Level	Passive Hammer	Bipolar	Opportunistic
Freehand	0.02	0.18	0.37
Passive Hammer		0.15	1.00
Bipolar			0.62

Table S20. p-Values from pairwise comparisons of percentage blank mass into flakes using Wilcoxon rank-sum tests

Naivety Level	Passive Hammer	Bipolar	Opportunistic
Freehand	0.45	0.69	1.00
Passive Hammer		0.04	1.00
Bipolar			0.46

Table S21. p-Values from pairwise comparisons of proportion of successful debitage from total debitage using Wilcoxon rank-sum tests

Naivety Level	Passive Hammer	Bipolar	Opportunistic
Freehand	0.05	0.04	0.35
Passive Hammer		0.40	0.63
Bipolar			0.32

Table S22. Eigenvalues from PCA

Number	**Eigenvalue**	**Percent**		**Cum. Percent**
1	4.2478	38.616		38.616
2	2.5155	22.868		61.484
3	1.7465	15.877		77.361
4	0.8072	7.338		84.700
5	0.7605	6.914		91.614
6	0.3960	3.600		95.214
7	0.2713	2.467		97.681
8	0.1625	1.477		99.157
9	0.0653	0.594		99.751
10	0.0219	0.199		99.951
11	0.0054	0.049		100.000

Table S23. Loading matrix for the PCA (showing only the first six principal components).

	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6
%UnifacialCores	0.62661	-0.41045	0.32396	-0.39755	-0.14129	0.12960
%bipolarcores	0.64956	-0.42804	0.38868	0.42667	0.16556	-0.03396
%cores	0.84705	0.29131	-0.23646	0.04168	0.17414	0.07224
%angularfrags	-0.55200	-0.05760	-0.28524	-0.09896	0.74988	0.14464
coresize	0.72211	0.49353	0.16426	-0.27005	0.19547	-0.28352
RatioFlaketoCore	-0.00231	0.30648	-0.81354	0.36485	-0.20926	-0.03860
FlakeScarCount	-0.28817	0.75924	0.53569	0.16463	0.00433	0.03590
RatioFlakeScarCoreSize	-0.38916	0.71204	0.52179	0.20288	0.03595	0.07484
FlakeSize	0.80056	0.48677	-0.23617	-0.04861	0.09794	-0.20319
FlakeThickness	0.69372	0.47342	-0.15161	-0.08540	-0.08734	0.46754
Percussion	0.73002	-0.44001	0.16239	0.41427	0.14139	0.05744

Table S24. Tool use data based on the behavior coding.

Individual	*N* of AS	Cutting Attempts		Cutting Successes		Cutting attempts DB1	Cutting attempts DB1 : cutting attempts	FT/CT Preference
		N	%	*N*	%			
P1	73	35	47.9	33	94.3	33	0.94	FT
P2	36	8	22.2	1	12.5	8	1.00	FT
P3	55	55	100.0	54	98.2	46	0.84	FT
P4	69	56	81.2	47	83.9	30	0.54	OP
P5	76	71	93.4	71	100.0	5	0.07	CT
P6	77	4	5.2	3	75.0	2	0.50	OP
P7	81	65	80.2	53	81.6	32	0.49	OP
P8	74	73	98.6	70	95.9	35	0.48	OP
P9	63	50	79.4	49	98.0	31	0.62	FT
P10	81	32	39.5	29	90.6	18	0.56	OP
P11	70	58	82.9	54	93.1	35	0.60	FT
P12	80	73	91.2	71	97.3	30	0.41	OP
P13	74	59	79.7	50	84.7	56	0.95	FT
P14	72	57	79.2	55	96.5	50	0.88	FT
P15	97	53	54.6	48	90.6	51	0.96	FT
P16	72	33	45.8	32	97.0	24	0.73	FT
P17	82	68	82.9	54	79.4	33	0.48	OP
P19	90	60	66.7	54	90.0	32	0.53	OP
P21	137	8	5.8	2	25.0	4	0.50	OP
P22	45	39	86.7	26	66.7	34	0.87	FT
P23	210	1	0.5	1	100.0	0	0.00	CT
P24	92	59	64.1	54	91.5	21	0.36	CT
P25	96	64	66.7	55	85.9	46	0.72	FT
P26	80	61	76.3	54	88.5	14	0.23	CT
P27	58	64	94.8	55	98.2	34	0.62	FT
P28	65	65	100.0	64	98.5	64	0.98	FT
P29	103	52	50.5	52	100.0	0	0.00	CT
P30	84	67	79.8	67	100.0	0	0.00	CT
Overall	2299	1381	57.0	1257	95.6	768	0.56	n.a.
Mean	81.9	49.3	n.a.	44.9	n.a.	27.4	n.a.	n.a.
Std. dev.	31.4	21.4	n.a.	21.2	n.a.	18.2	n.a.	n.a.

Appendix III: Study documents

Explanation of candidate contribution to collaborative work

This chapter is based on collaborative work. The text contains (modified) excerpts from the main and supplementary text of the following publication:

Snyder, W.D., Reeves, J.S., & Tennie, C. (2022). Early knapping techniques do not necessitate cultural transmission. *Science Advances, 8*(27), eabo2894. doi:10.1126/sciadv.abo2894

Conceptualization: C.T., W.D.S., J.S.R. Investigation: W.D.S. Visualization: W.D.S. Formal analysis: W.D.S., C.T. Writing (original draft): W.D.S. Writing (review and editing): W.D.S., C.T.

Questionnaire on the Prior Experiences of Study Participants

All participants were given the Questionnaire on the Prior Experience of Study Participants. This questionnaire was given to the participants after their testing was complete in order to not compromise the integrity of the study and the naivety of the individuals (*pre*-test questionnaires would have compromised these). Responses to this questionnaire were used to determine the naivety level of each participant.

Fragebogen zu den Vorerfahrungen der Studienteilnehmer Datum: _________ Teilnehmer ID: __________ Hinweis: Ihre Antworten, egal wie diese ausfallen, werden keine negativen Folgen für Sie haben. Sie können sämtliche monetären Belohnungen („Preise") sowie sämtliche Aufwandsentschädigungen, die Sie bisher bekommen haben, behalten. Bitte beantworten Sie folgende Fragen möglichst knapp, oder kreisen (falls diese Optionen vorhanden ist) die passende Antwort ein. 1. Sind Sie derzeit an einer Universität eingeschrieben? Falls ja, für welchen Studiengang? 2. In welchem Semester befinden Sie sich? 3. Falls Sie kein Studierender sind, was ist Ihr höchster Schulabschluss? Wenn Sie zu einem früheren Zeitpunkt an einer Universität eingeschrieben waren, bis zu welchem Semester hatten Sie studiert? 4. Haben Sie jemals an Kursen teilgenommen , in denen Steinwerkzeuge zumindest *erwähnt* wurden? O JA O NEIN a. Falls ja, beschreiben Sie bitte entsprechende Kurse (und den Veranstaltungsort). Gehen Sie bitte insbesondere die Art und Tiefe der Informationen, die dort über Steinwerkzeuge vermittelt wurden ein, v.a. bezüglich der Art und Weise der Herstellung von Steinwerkzeugen. b. Falls ja, wann haben Sie an diesem Kurs/diesen Kursen teilgenommen (z.B. „vor dieser Studie, im Jahre XYZ")? [Mehrfachnennung möglich]? 5. Haben Sie bereits an *Studien* zu Steinwerkzeugen (z.B. zu deren Beschreibung oder Herstellung) teilgenommen? O JA O NEIN a. Falls ja, bitte die Studie(n) im Folgenden kurz beschreiben (und Studien-Ort bitte angeben). [Mehrfachnennung möglich]	*Questionnaire on the Prior Experience of Study Participants* Date: _________ Participant ID: __________ Note: Your answers, regardless of how they turn out, will lead to no negative consequences for you. You can keep all monetary rewards (prizes), along with all the hourly compensation, that you have thus far received. Please answer the following questions as succinctly as possible, or cross out the suitable answer (if this option is provided). 1. Are you currently enrolled at a university? If yes, in which degree of study? If no, please continue with question 3. 2. In what semester are you currently? 3. If you are not a student, what is your highest attained degree? If you were enrolled at a university at an earlier point in time, up to what semester had you studied? What did you study? 4. Did you take part in a course, in which stone tools were at minimum mentioned? O YES O NO a. If yes, please describe the corresponding course(s) (and the course setting). Please provide especially the type and depth of information about stone tools that was conveyed there, above all as to the ways of stone tool production. b. If yes, when do you take part in this course (e.g., "Before this study, in the year XYZ")? [Multiple answers possible]? 5. Have you already participated in studies on stone tools (e.g., on their description or their manufacture)? O YES O NO a. If yes, please briefly describe them (and provide the setting/study site). [Multple answers possible]

b. Falls ja, wann haben Sie an dieser Studie/diesen Studien teilgenommen (z.B. „vor dieser Studie, im Jahre XYZ“)? [Mehrfachnennung möglich]?	b. If yes, when did you take part in this study/these studies (e.g., “Before this study, in the year XYZ”)? [Multiple answers possible]?
6. Haben Sie schon einmal an einem/mehreren praktischen Kursen teilgenommen, deren Ziel die Herstellung von Steinwerkzeugen war? O JA O NEIN	6. Have you already taken part in one or more practical courses, in which the goal was the production/manufacture of stone tools? O YES O NO
a. Falls ja, beschreiben Sie bitte kurz entsprechende Kurse (und den Veranstaltungsort).	a. If yes, please briefly describe them (and the course setting) [Multiple answers possible]
b. Falls ja, wann haben Sie an diesem Kurs/diesen Kursen teilgenommen (z.B. „vor dieser Studie, im Jahre XYZ“)? [Mehrfachnennung möglich]	b. If yes, when did you take part in this course/these courses (e.g., “before this study, in the year XYZ”)? [Multiple answers possible]
7. Haben sie sich jemals aktiv über Steinwerkzeuge informiert oder passiv etwas über Steinwerkzeuge gelernt (z.B. durch Bücher, Radio, Fernsehen, Internet, Museumsbesuch etc.)? O JA O NEIN	7. Have you ever actively informed yourself about stone tools or passively learned about stone tools (e.g., via books, radio, television, internet, museum visits, etc.)? O YES O NO
a. Falls ja, bitte geben Sie an, welche Art des Medium (z.B. Radio). [Mehrfachnennung möglich]	a. If yes, please indicate the kind of medium (e.g., radio). [Multiple answers possible]
b. Falls ja, wie detailliert waren diese Informationen (z.B. Zeitleiste der Steinwerkzeuge; Art und Weise der Herstellung etc.)? [Mehrfachnennung möglich]	b. If yes, how detailed was this information (e.g., a timeline of stone tool types; ways of stone tool production, etc.)? [Multiple answers possible]
c. Falls ja, wann haben Sie diese Information passiv oder aktiv erhalten? (z.B. „vor dieser Studie, im Jahre XYZ“)? [Mehrfachnennung möglich]	c. If yes, when did you receive this information? (e.g., „Before this study, in the year XYZ“)? [Multiple answers possible]
8. Haben Sie schon einmal Abbildungen (Illustrationen, Fotografien, Videos, Diagramme, Museumsaustellungen, etc.) die frühe Steinwerkzeuge zeigen, gesehen? O JA O NEIN	8. Have you already seen depictions of early stone tools (illustrations, photographs, videos, diagrams, museum exhibitions, etc.)? O YES O NO
a. Falls ja, bitte beschreiben Sie kurz (und, wenn Sie wollen, zeichnen) Sie, was Sie gesehen haben. [Mehrfachnennung möglich]	a. If yes, describe briefly describe (and, if you want, draw) what you have seen. [Multiple answers possible]
b. Falls ja, wann haben Sie diese zum ersten Mal gesehen (z.B. „vor dieser Studie, im Jahre XYZ“)? [Mehrfachnennung möglich]	b. If yes, when did you see this/these for the first time (e.g., “Before this study, in the year XYZ”)? [Multiple answers possible]
9. Haben Sie sich jemals den *Prozess* der Steinwerkzeugherstellung angesehen, also die Art und Weise der Herstellung (einschließlich live, YouTube – Videos, Dokumentarfilmen etc. – oder auch nur als Illustrationen (Bilder oder Video))? O JA O NEIN	9. Have you ever viewed the process of stone tool manufacture, like the manner of production (including live, YouTube videos, documentaries, etc. – also as just illustrations (pictures or videos))? O YES O NO
a. Falls ja, bitte beschreiben Sie, was Sie gesehen haben (v.a. Detailgrad der Werkzeugherstellung) und wo. [Mehrfachnennung möglich]	a. If yes, please describe what you have seen (above all, the level of detail of the stone tool production) and where. [Multiple answers possible]
b. Falls ja, wann haben Sie dies zum ersten Mal gesehen (z.B. „vor dieser Studie, im Jahre XYZ“)? [Mehrfachnennung möglich] 10. Haben Sie jemals Anleitungen zur Steinwerkzeugherstellung bekommen?	b. If yes, when do you see this for the first time (e.g., “Before this study, in the year XYZ”)? [Multiple answers possible]

O JA O NEIN

a. Falls ja, bitte beschreiben Sie, welche Art von Anleitung dies war (v.a. Detailgrad der Werkzeugherstellung) und wo Sie diese erhalten haben. [Mehrfachnennung möglich]

b. Falls ja, wann haben Sie zum ersten Mal diese Anleitungen erhalten (z.B. „vor dieser Studie, im Jahre XYZ“)? [Mehrfachnennung möglich]

11. Würden Sie von sich sagen, dass Sie in generellem Bezug auf Steinwerkzeuge generell gute Kenntnisse haben – in dem Sinne, dass Sie diese Kenntnisse außerhalb unserer Studie gewonnen haben?
O JA O NEIN

12. Würden Sie von sich sagen, dass Sie in Bezug auf die *Herstellung* von Steinwerkzeuge gute Kenntnisse haben – in dem Sinne, dass Sie diese Kenntnisse außerhalb unserer Studie gewonnen haben?
O JA O NEIN

13. Haben Sie das Gefühl, das Sie vor der Studie relevante Informationen zu Steinwerkzeugen und/oder zur Steinwerkzeugenherstellung besaßen, die durch die obigen Fragen nicht abgedeckt sind?
O JA O NEIN

a. Falls ja, bitte beschreiben Sie, welche Art von Informationen (v.a. Detailgrad der Werkzeugherstellung) und wo Sie diese erhalten haben. [Mehrfachnennung möglich]

b. Falls ja, wann haben Sie diese Informationen erhalten (z.B. „vor dieser Studie, im Jahre XYZ“)? [Mehrfachnennung möglich]

14. Haben Ihnen irgendwelche Information, die Sie vor der Studie erhalten haben, in dieser Studie geholfen Steinwerkzeuge herzustellen?
O JA O NEIN

a. Falls ja, bitte beschreiben Sie, welche Art von Informationen (v.a. wie viel Detail zur Werkzeugherstellung) und wo Sie diese erhalten haben. [Mehrfachnennung möglich]
b.
c. Falls ja, wann haben Sie diese Informationen erhalten (z.B. „vor dieser Studie, im Jahre XYZ“)? [Mehrfachnennung möglich]

15. Gibt es weitere Informationen zu Ihrer Person , die Sie eventuell für relevant für uns/Steinwerkzeuge betrachten?
O JA O NEIN

a. Falls ja, bitte unten beschreiben:

10. Have you ever received instructions of stone tool making?
O YES O NO

a. If yes, please describe what kind of instruction this was (above all, the level of detail of the stone tool production) and where you received it. [Multiple answers possible]

b. If yes, when did you receive this instruction (e.g., “Before this study, in the year XYZ”)? [Multiple answers possible]

11. Would you yourself say, that you generally have good knowledge in relation to stone tools – in the sense that you earned this knowledge outside of our study?
O YES O NO

12. Would you yourself say, that you have good knowledge in relation tot he specific means of production of stone tools – in the sense that you earned this knowledge outside of our study?
O YES O NO

13. Do you have the feeling that you possessed or received relevant information to stone tools or stone tool production before or during the study that was not covered by the questions above?
O YES O NO

a. If yes, please describe what kind of information (above all, the level of detail of stone tool production) and where you received this. [Multiple answers possible]

b. If yes, when did you receive this information (e.g., “Before this study, in the year XYZ)? [Multiple answers possible]

14. Did you receive any information before or during this study, which helped you produce stone tools in this study?
O YES O NO

a. If yes, please describe what kind of information (above all, how much detail of stone tool production) and where you received this information. [Multiple answers possible]

b. If yes, when did you receive this information (e.g., “Before this study, in the year XYZ”)? [Multiple answers possible]

15. If there any more information about you that you potentially consider relevant for us/stone tools?
O YES O NO

a. If yes, please describe below:

Post-study Questionnaire (for participants who were not successful)

This questionnaire was given only to participants who, during the entire test, did not succeed at the task in the sense of creating and using cutting tools.

Post-Studie Fragebogen (von Teilnehmern, die nicht erfolgreich waren) *Teil 1* Datum: _________ Teilnehmer ID: __________ Bitte beantworten Sie die folgenden Fragen. Hinweis: Ihre Antworten, egal wie diese ausfallen, werden keine negativen Folgen für Sie haben. 1. Was war Ihrer Meinung nach das Ziel unseres Experiments – was, glauben Sie, haben wir versucht zu untersuchen? 2. Was glauben Sie, hat Sie vom erfolgreichen Lösen der Aufgabe abgehalten? Beschreiben Sie bitte kurz die Faktoren, die Sie dabei womöglich gehindert haben.	*Post-Study Questionnaire (for Participants who were not successful)* *Part 1* Date: ________ Participant ID: _________ Please answer the following questions. Notice: Your answers, regardless of how they turn out, will not result in negative consequences for you. 1. What, in your opinion, was the purpose of our experiment – what, in your belief, did we attempt to research? 2. What do you think prevented you from a successful solution of the task? Please briefly describe the factors that have possibly hindered you from doing so.
Post-Studie Fragebogen (von Teilnehmern, die nicht erfolgreich waren) *Teil 2* *(Die Teilnehmer bekommen zuvor eine Kopie von der „Nachbesprechung der Problemlösestudie“.)* Datum: _________ Teilnehmer ID: __________ Beantworten Sie bitte kurz die folgenden Fragen und denken Sie an die Information, die Sie gerade erhalten haben („Nachbesprechung der Problemlösestudie“). 1. Denken Sie, dass Sie das Ziel des Experiments vollständig, oder nahezu vollständig, verstehen? O JA O NEIN 2. Glauben Sie, im Hinblick auf die eben erhaltenen Informationen, dass noch andere Faktoren Sie vom erfolgreichen Lösen der Aufgabe abgehalten haben? Beschreiben Sie bitte kurz die Faktoren, die Sie dabei vermutlich besonders gehindert haben.	*Post-Study Questionnaire (for Participants who were not successful)* *Part 2* *(The participant receives beforehand a copy of the “Problem-solving Study Debriefing”)* Date: ________ Participant ID: _________ Please briefly answer the following questions and think about the information that you have just received (“Problem-solving Study Debriefing”). 1. Do you think that you understand the purpose of the experiment completely or nearly completely? O YES O NO 2. Do you think, thinking back on the information you just received, that yet more factors prevented you from successful solution of the task? Please briefly describe the factors that have probably hindered you from doing so.

Additional questions in case of a partial solution

This questionnaire was given only to participants who, during the entire test, did not fully succeed at the task in the sense that, while they created cutting edge on glass and/or stone, they never *used* the created cutting edge to overcome the rope of the tendon box.

Weitere Fragen im Falle von einer Teillösung der Aufgabe Datum: _________ Teilnehmer ID: __________	*Additional questions in case of a partial solution* Date: _________ Participant ID: __________
Bitte beantworten Sie die folgenden Fragen. Hinweis: Ihre Antworten, egal wie diese ausfallen, werden keine negativen Folgen für Sie haben.	Please answer the following questions. Note: Your questions, regardless of how they turn out, will lead to no negative consequences for you.
Das Ziel dieses Experiments, also die vollständige Lösung der Aufgabe, beinhaltete a) die Erschaffung von Objekten mit scharfen Kanten, alleinig durch entsprechende Verwendung der bereitgestellten Materialien, und b) das Nutzen dieser scharfen Kanten, um Zugang zu den Belohnungen zu erhalten (also durch Durchschneiden von einem Seil oder einer Membran).	The goal of this experiment was twofold a) the creation of sharp edges via the usage of the provided materials, and b) the cutting of either the rope or the membrane of the apparatuses, in order to receive the reward(s).
Sie bekommen diesen Fragebogen, weil Sie zwar ein Objekt (oder mehrere Objekte) mit solchen scharfen Kanten erschaffen haben, aber diese Kanten dann nicht genutzt haben, um an die Belohnungen zu gelangen.	You have created one or more objects with sharp edges, but you did not see the task all the way through to the finish.
1. Haben Sie während des Tests erkannt, dass Sie scharfe Kanten erschaffen haben? O JA O NEIN a. Falls ja, warum haben Sie (Ihrer Meinung nach) diese scharfe Kanten nicht verwendet, um mindestens einen von den zwei Apparaten zu öffnen und die enthaltene Belohnung zu bekommen?	1. Did you recognize that you had created such an objects or objects with sharp edges? O YES O NO a. If yes, why did you not use such sharp edges, in order to open one of the apparatuses and receive the reward?
2. Was glauben Sie, hat Sie vom erfolgreichen Lösen der Aufgabe abgehalten? Beschreiben Sie bitte kurz die Faktoren, die Sie dabei womöglich gehindert haben.	2. What do you believe prevented you from a successful solution of the task? Please briefly describe the factors that possible would have stopped you.
Vielen Dank für das Ausfüllen dieses Fragebogens. Bitte denken Sie daran, alle Informationen zur Studie vertraulich	Thank you for filling out this questionnaire. Please bear in mind to handle all information about this study in

zu behandeln und diese nicht anderen Personen weiterzugeben. Falls Sie weitere Fragen in Bezug auf das Experiment stellen möchten, kann der Versuchsleiter diese nun gerne beantworten. Falls Sie später weitere Fragen haben, können Sie diese an William Snyder schicken. Vielen Dank für Ihre Teilnahme!	confidentiality and to not pass on this information to any other person. If you have further questions in relation to the experiment that you wish to pose, the lead experimenter can gladly answer those now. If you have more questions later, you can send these to William Snyder. Thank you for your participation!

Experimenter script

In order to maintain the same method for each participant session, the experimenter followed a standardized script for each test, to prevent experimenter error and to prevent variation in information given across participants. Icons were included to make for quick and easy accessibility of the relevant responses to specific situations. Thus, the script is presented here in the same exact form as the printout used by the experimenter.

Key phrases that need to be standardized to ensure that all participants get the same info:

Pre-test preparations:

Understanding of pre-study documents, incl. participant information and consent

- Haben Sie alle Dokumente, die Sie von mir per Email bekommen haben, komplett gelesen und verstanden? Sie können mir jetzt Fragen stellen, die Sie noch über die Dokumente haben.
    - EN: Have you completely read and understood all documents, that you received before by Email? Please pose any questions you have about the documents now.

If *no* to previous question:

- Bevor Sie teilnehmen dürfen, müssen Sie alle Dokumente durchlesen, und Ihnen durch Unterschrift zustimmen.
    - EN: Before you may participate, you must now read through all documents, in order to know and clearly understand your rights and the requirements of this study.

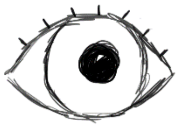
Security measures

- Bevor Sie in den Testraum hineinkommen dürfen, müssen Sie sich zuerst (jedes Mal!) diese Schutzausrüstung anziehen. Sie müssen alle Teile anziehen. Wenn diese Ausrüstung während des Tests unangenehm wird oder irgendwas wehtut, informieren Sie mich bitte. Bevor Sie die Schutzausrüstung ausziehen dürfen, müssen Sie erst den Testraum verlassen. Durch die Schutzausrüstung ist es zwar sehr unwharscheinlich, dass Sie sich verletzen, aber falls Sie sich in irgendeiner Weise trotzdem verletzen sollten, informieren Sie mich bitte sofort.
    - EN: Before you may enter the testing space, you must first (and every time) put on the safety gear. You must put on every piece of this gear. If you are uncomfortable during the experiment or something hurts, please inform me immediately. Before you can remove the safety gear, you must first leave the testing space. It is also important that you inform me of injuries of any kind (within the session).
- Falls möglich, sollten Sie Schmuckstücke von Ihrer Händen abziehen. Ich bitte Sie auch, mir ihr Handy zu geben. Aber keine Angst: Ihr Handy und ihr Schmuck werden auf dieses Bord in eine Box gelegt und bleiben dort – für Sie sichtbar - bis zum Ende der Sitzung. Sie dürfen Ihr Handy während der Testsitzung nicht nutzen (ich empfehle, es nun lautlos zu schalten). Dies ist nötig, weil wir Sie auf Ihre eigene Lösungsansätze testen, und deswegen müssen wir verhindern, dass Sie nach Lösungsansätzen im Internet suchen.

    - EN: If possible, you should remove any jewelry from your hands. I also ask that you give me your phone. But no worries: Your phone and jewelry will be placed on this shelf in a box – visible to you - and will stay there until the end of the session. You may not use your phone during the session (I recommend to set it to silent now); we are testing for your own individual solving approaches, and therefore, we want to prevent that you search for outside (=Internet) solutions.

Beginning:

[Before continuing, press record on all cameras and clearly show the coding sheet to each].



- Sobald ich "Start" sage, dürfen Sie anfangen. Wenn ich „Stopp" sage, halten Sie bitte sofort inne (was auch immer Sie in dem Moment tun, müssen Sie dann stoppen).
    - EN: When I say Start, you may begin. When I say Stop, you must immediately stop (whatever you are doing, you must absolutely stop)
- Sehen Sie hier diese Aufgabe [point to the apparatus]? Bei *dieser* Aufgabe [tendon box] hindert Sie *dieser* Teil [point to rope] daran *diese* Tür zu öffnen um an die Belohnung zu gelangen. **Beachten Sie bitte die *anderen* Teile der Aufgabe *gar nicht*. Mit anderen Worten: konzentrieren Sie sich ausschließlich auf diesen [point] Teil der Aufgabe.** Zu Ihrer Information: Es *ist* möglich, diese Aufgabe mit Hilfe der Ihnen zur Verfügung gestellten Objekte zu lösen.
    - EN: Do you see this task here? With this task, this part [the rope] prevents you from opening this door and accessing the reward. Do not pay any attention to the parts of the apparatus. In other words: concentrate exclusively on this part of the task. For your information: it is possible to solve this task with the help of the objects that are placed for your availability.
    - Possible replies:
        - Meinen Sie diese Schnur/diese Seile/usw.?
            - EN: Do you mean this string/rope/ etc.?
            - Confirm, but do not repeat the term they use (refer to them always as "Teil")
- Sie können sich jetzt gerne die Aufgabe anschauen. Wollen sie? [if yes, allow. If no, continue]
    - EN: You can now look at the apparatus. Would you to do so?
- Sie können alle diese Objekte [point to everything, except the apparatus] anheben und aufnehmen. Die Aufgabe selbst müssen Sie allerdings stehenlassen.
    - EN: You can lift and pick up all of these objects. You must leave the apparatus itself as it is.
- Sie können *alle* diese Objekte [point slowly at granite, hammerstones and glass blank] benutzen um die Aufgabe zu lösen, in jeglicher Weise - und ganz wie Sie es für richtig halten.

    - EN: You are allowed to lift all objects in the testing space except for the apparatus. You may use all of these objects in any fashion and as you see fit in order to solve the task.
- Es gibt *mehrere* Methoden, mit denen man die Aufgabe lösen kann. Einige Methoden dürfen wiederholt angewendet werden; andere dürfen nicht wiederholt werden. Ich werde Sie jeweils informieren, falls eine Methode *nicht* wiederholbar ist.
    - EN: There are several methods, with which one can solve the task. Some methods may be repeatedly used; others cannot be repeated. I will inform you each time, if a method is not repeatable.
- Sehen Sie die Belohnungen / das Monopoly-Geld? (point to apparatus and wait for an affirmative)
    - EN: Do you see the rewards? (point to apparatus and wait for an affirmative)
- Versuchen Sie, die Aufgabe zu öffnen, um die jeweils enthaltene Belohnung zu bekommen – diese Belohnung wird Ihnen am Ende der Sitzung in Echtgeld ausbezahlt (das heißt, *zusätzlich* zu ihrem generellen Aufwandsentschädigung pro Stunde)
    - EN: Attempt to open this apparatus, in order to be paid for the reward inside – this will be paid out to you at the end of the session in real money (in addition to the base hourly rate).
- Um die Videoaufnahmen zu synchronisieren werde ich nun kurz einen einzelnen Lichtblitz auslösen. Ich empfehle Ihnen kurz die Augen zu schliessen oder wegzuschauen. Auf drei löse ich den Blitz aus. 1, 2 , 3.
    - EN: In order to synchronize the video recordings, I will trigger shortly a single photo flash. I recommend that you briefly close your eyes or look away. On three I will trigger the flash. 1, 2, 3.
- Start, oder: Sie dürfen gerne anfangen. [Experimenter starts stopwatch and timers]
    - EN: Start, or: You may begin.

During Testing:

If they ask at any point what they can and cannot do:

- Sie können mich zwar generelle Fragen stellen aber ich empfehle Ihnen das nicht zu tun. Ich darf Ihnen sowieso keine Lösungen nennen und die Antworten könnten Sie zudem verwirren. Deshalb: Probieren Sie am Besten einfach aus, was Ihnen einfällt. Ich würde Sie dann schon stoppen.
    - EN: You can ask me general questions but I recommend that you don't. I am not allowed to name any solutions, and the answers could confuse you. As such: simply try out, whatever crosses your mind. I would stop you [if it is not appropriate].
- Kann ich X, Y, Z tun?
    - EN: Can I do X, Y, Z?
    - Sie können alles versuchen, mit Ausnahme von allen jenen Methoden bei denen ich dazusagte dass Sie sie nicht mehr machen dürfen (also: XXX [list the ones that apply at that moment, starting with „Ihnen sind bisher noch keine Methoden verboten worden“]).

        - EN: You can use any method, with the exception of all methods that I said to you that they are no longer allowed to be done. (also: XXX [list the ones that apply at that moment, starting with "For you, there are so far no methods that have been forbidden"

- Kann ich/darf ich X [the stones, the core, anything other than the boxes of the apparatus] zerbrechen?
    - EN: Can I/am I allowed to break X?
    - Probieren Sie einfach das aus, was Ihnen einfällt. Ich würde Sie dann schon stoppen.
        - EN: Simply try out, whatever comes to mind. I would stop you [if it is not appropriate].

If they insist:

- Alles was ich Ihnen sagen kann ist: Sie können alle Objekte benutzen die vor Ihnen liegen, um die Aufgabe zu lösen.
    - EN: You can use all objects that are lying in front of you to solve the apparatus.

If they are vocally angry, frustrated, confused, etc.:

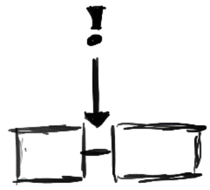
- Der Versuchsaufbau erlaubt mir nicht Ihnen Tips oder Anleitungen zu geben.
    - The experimental setup does not allow me to give you tips or instructions.
- If they are frustrated about changes: Bitte erinnern Sie sich daran, dass die Teilnehmerinformation folgenden Satz enthielt: „Der Versuchsleiter behält sich vor, die Regeln während des Versuchs zu vervollständigen oder abzuändern, um einen ordnungsgemäßen Versuchsablauf zu sichern."
    - EN: Please also recall that the Participant Information sheet contains the following phrase: „The investigator reserves the right to amend or modify the rules during the trial in order to ensure proper experimental procedure and protocol."

Additionally, offer general encouragement when the participant is struggling to dissuade loss of motivation.

- [zum Beispiel]: Sie können gerne weitermachen. Sie haben noch Zeit. Machen Sie daher ruhig weiter, und probieren Sie einfach.
    - EN: You can gladly continue. There is still time. You still have time. Steadily continue and simply try things out.

If they engage in dangerous behaviour or attempt to lift the anvil above the ground:

- Stopp. Machen Sie bitte XXX nicht [whatever halts their dangerous actions]- Eine Wiederholung davon wird zu der Beendung dieser Sitzung und keiner weiteren Teilnahme von Ihnen führen.
    - Stop. Do not do XXX. Repetition thereof will lead to the ending of this session and no further involvement of you in the study.
- Stop. Bitte dieses Objekt nicht anheben.



- Stop. Please do not lift up this object [granite anvil].

If they continue [more than 5 minutes in total] to explore the outside of the apparatus with their hands or objects (including but not limited to: touching and trying to turns screws, touching and hitting the non-solution-related parts of the apparatuses, trying to pull on the door of the tendon box, etc.):

- Stop. Bitte halten Sie kurz inne. Ich möchte Sie noch einmal erinnern [point to the apparatus and repeat the description].
  - EN: Stop. Please cease briefly. I would like to remind you once more the description of the apparatuses.

Repeat as necessary:

- Bei dieser Aufgabe [tendon box] hindert Sie dieser Teil [point to rope] daran diese Tür zu öffnen um an die Belohnung zu gelangen. Beachten Sie nicht den Rest der Aufgabe gar nicht. Zu Ihrer Information: Es *ist* möglich, diese Aufgabe mit den Ihnen zur Verfügung gestellten Objekten zu lösen.
  - EN: With this task, this part prevents you from opening the door and accessing the reward. Do not pay any attention to the rest [of the apparatus]. For your information: it is possible to solve these tasks with the objected that are placed for your availability.

If they use a method besides cutting (after they complete the method once):

If they rub on the rope:

- Ab jetzt dürfen Sie nicht mehr die Aufgabe durch Reiben lösen. Reiben ist nun keine erlaubte Lösung mehr, denn das nicht der Lösungsansatz an dem wir interessiert sind. Versuchen Sie also weiterhin eine Lösung zu finden.
  - EN: Henceforth, you may no longer solve the task by rubbing. Rubbing is no longer an allowed solution, as it is not the solution that we are interested in. Please continue to attempt to find another solution. [Note: Here and for other 'unwanted' solutions, we include this specific phrasing due to Pilot Participant 2 mistakenly interpreting the goal of the experiment to be diversity of solutions and therefore potentially self-directing their behavior to accomplish this.]

If they push down or up with a blunt surface or edge:

- Ab jetzt dürfen Sie nicht mehr die Aufgabe durch Pressen oder Ziehen lösen. Pressen oder Ziehen sind nun beides keine erlaubte Lösung mehr, denn das nicht der Lösungsansatz an dem wir interessiert sind. Versuchen Sie also weiterhin eine Lösung zu finden.
  - EN: Henceforth, you may no longer solve the task by pressing down or pulling up. Pressing is no longer an allowed solution, as it is not the solution that we are interested in. Please continue to attempt to find another solution.

If they use their hands to pull or push open an apparatus:

- Ab jetzt dürfen Sie die Aufgabe nicht mehr ohne Zurhilfenahme von Objekten lösen. Die Lösung ohne Objekte ist nun keine erlaubte Lösung mehr, denn das nicht der Lösungsansatz an dem wir interessiert sind. Versuchen Sie also weiterhin eine Lösung zu finden.
  - EN: Henceforth, you may no longer solve the task without the help of the objects. The solution without objects is no longer an allowed solution, as it is not the solution that we are interested in. Please continue to attempt to find another solution.

If they open the apparatus by knocking together two objects:

- Ab jetzt dürfen Sie die Aufgabe nicht mehr durch Aufklopfen lösen. Aufklopfen ist nun keine erlaubte Lösung mehr, denn das nicht der Lösungsansatz an dem wir interessiert sind. Versuchen Sie also weiterhin eine Lösung zu finden.
  - EN: Henceforth, you may no longer solve the task by knocking together two objects. This knocking is no longer an allowed solution, as it is not the solution that we are interested in. Please continue to attempt to find another solution.

If after two hours they are not successful and are to receive the Impossible Flake:

- Bitte nehmen Sie dieses Objekt [the impossible flake]. Sie dürfen es einmalig benutzen, um die Aufgabe zu lösen. Danach wird es Ihnen wieder weggenommen.
  - EN: Take this thing/object. You may use it once to solve the task. Afterward, it will be taken away.

When they succeed at cutting:

- Stopp! Halten Sie bitte inne. Danke. Bitte legen Sie nun das Ding in Ihrer Hand auf den Boden. Ich benötige nun ein wenig Zeit, um Daten aufzunehmen. Setzen Sie sich bitte derweil auf den Stuhl dort in der Ecke, bis ich Ihnen sage das sie zurückkehren können.
  - EN: Stop! Please cease [what you are doing]. Thank you. Lay the thing in your hand [=flake or core-tool] on the ground. I require a little time to record data. Sit down please meanwhile on the chair in the corner, until I say that you can return.

First time: explain the rule that successful tools are to be transferred to the experimenter including reference to statement of new rules in consent form

- DE: Bitte erinnern Sie sich, dass die Teilnehmerinformation folgenden Satz enthielt: „Der Versuchsleiter behält sich vor, die Regeln während des Versuchs zu vervollständigen oder abzuändern, um einen ordnungsgemäßen Versuchsablauf zu sichern.“ Ab jetzt gilt nun die folgende Regel: Jedes Mal, wenn Sie die Testaufgabe lösen, müssen Sie das Ding, mit dem Sie die Aufgabe gelöst haben, mir übergeben.
  - EN: Please recall that the Participant Information sheet contains the following phrase: „The investigator reserves the right to amend or modify the rules during the trial in order to ensure proper experimental procedure and

protocol.” From now on, the following rule applies: every time that you solve the task, you must hand over to me the thing with which you have solved it.

If they shatter the core producing multiple pieces:

- Stopp! Halten Sie bitte inne. Bevor Sie weitermachen dürfen, müssen Sie nun ein Teilstück wählen, mit dem Sie weitermachen wollen. Alle anderen Teile werden von mir eingezogen.
    - EN: Stop! Please pause. Before you are allowed to continue, you must now select one piece with which you want to continue working. All other pieces will be confiscated by me.

Upon creation of a core-tool and use of edge on core-tool to cut apparatus:

- Stopp! Halten Sie bitte inne. Bitte bleiben Sie genau so wie Sie gerade sind. Ich werde nun mit einem Stift den Bereich markieren, den Sie gerade genutzt haben, um die Aufgabe zu öffnen. Bitte beachten Sie die neue Regel: Markierte Bereiche dürfen Sie nicht mehr nutzen um den Aufgabe zu öffnen.
    - EN: Stop. Please cease what you are doing. Please stay exactly as you are now. I will mark the edge that you just used to solve the apparatus with a pen. Please be aware of the new rule: marked edges may no longer be used to open the apparatus.

After the first success, extra safety instruction due to sharp materials on ground:

- Für Ihre eigene Sicherheit knien Sie bitte von nun an nicht mehr auf dem Boden. Auf dem Boden können sich Splitter befinden.
    - EN: For your own safety, please do not kneel on the ground from now on. Sharp chips can be found on the ground.

After the first success:

- Dieser Ausdruck repräsentiert das von Ihnen gerade zusätzlich gewonnene Geld, also ihre Belohnung. Jedes Mal, wenn Sie die Aufgabe lösen, wird die jeweilig enthaltene Belohnung dort [in einer Box] gesammelt. Am Ende der Sitzung wird der Wert von allen Belohnungen zusammengerechnet. Sie werden dann die Gesamtsumme in Echtgeld erhalten (das heißt, sie bekommen sowohl die stündlichen Aufwandsentschädigungen als auch die gesammelten Belohnungen).
    - EN: This paper printout represents the money you have additionally won, otherwise your reward. Every time that you solve the task, the reward received thus far will be collected here. At the end of the session, the value of the rewards all together will be calculated. You will be paid the real money in one lump sum (that means, you receive both the hourly compensation and the rewards for successes).
- Bitte beachten Sie, dass ab jetzt die Ausdrücke nicht mehr lesbar ohne die Öffnung der Aufgabe werden (die Schriftseite wird nach unten gelegt). Die genaue Höhe der Belohnung entdecken Sie erst, wenn Sie die eder erfolgreich gelöst haben. Beachten Sie bitte: Die Höhe der Belohnungen wird ab jetzt durch eine Tombola [if they don‘t understand: „durch gemischte Karten“ or per „Lotterieverfahren“ or „zufällig“]

festgelegt. Die Zettel wurden vor der Sitzung von mir gemischt (d.H., randomisiert). Der Höchstwert der zukünftigen Belohnungen liegt bei 10 €.
    - EN: Be aware that from now on the paper printouts will no longer be legible without the opening of the apparatus (they will be placed face down). You will discover the value of the reward first when you successfully solve the task again. Please be aware: the values of the rewards will now be determined by a raffle/draw [if they don't understand: "by mixing of cards" or per lottery or "chance"]. The slips of paper were already shuffled by me before the session (i.e., randomized). The maximum value of future rewards is 10€.

If the participant breaks the hammerstone, including gradually or in one event, there is a protocol to be followed. 1/3 volume lost (as visually judged by experimenter) elicits a question to the participant. ½ volume lost means automatic replacement of hammerstone.

- Stopp. Ich kann Ihnen anbieten, dieses Objekt [point to hammer] gegen dieses [point to potential replacement] auszutauschen.. Möchten Sie es austauschen?
    - EN: Stop. I can offer you this: you can trade this object for this. Would you like to trade?
- Stopp. Dieses Objekt [point to current hammer] muss nun gegen dieses [point to new hammer] ausgetauscht werden.
    - EN: Stop. This object must now be traded out for this.

If they stop attempting to make flakes are or are unsuccessful at making a flake for an extended period of time (=min. of 5 minutes), they have already made flakes with this core, **and** more than an hour of testing has passed by. They cannot request a new core before they have intentionally made flakes on the core. Researcher's judgment of intentional flaking and core depletion is applied here.

- Hindert Sie etwas daran weiterzumachen?
    - EN: Is something preventing you from continuing?
- If they say yes [in the sense towards the problem of not being able to work the core further] and have made flakes: Sie können ein neues Objekt bekommen [hold up new core]. Sie können ein solches, neues Objekt alle 60 Minuten erhalten, falls Sie dann ein neues Objekt wünschen. Beachten Sie, falls Sie ein neues Object bekommen wollen, wird Ihnen jeweils das alte Objekt von mir abgenommen. Falls 60 Minuten seit der letzten Objektannahme noch nicht vergangen sind, müssen Sie warten bis diese 60 Minuten rum sind.
    - EN: You can receive a new object. You can receive such a new object every 60 minutes, in case you wish for one. Be aware, in case you do want to receive a new object, the old object will, in each case, be taken away by me. If 60

minutes have not yet passed since the last handout, you must wait until this 60 minutes is finished.

- [At the end of the study if they exhausted the core and want to end:] Sie befinden sich in der letzten Stunde der Sitzung. Falls sie nicht weitermachen können, und trotzdem ihre volle zeitliche Aufwansentschaedigung erhalten wollen, müssen Sie horbleiben und warten bis zum Ende der Testsitzung. Sie dürfen naturlich auch jederzeit gehen, aber Sie bekommen dann kein Geld für die restliche Zeit der Sitzung. Sie haben dann natürlich auch keine Möglichkeit mehr, weitere Belohnungen aus der Aufgabe zu erhalten.
    - EN: You are now in the last hour of the session. In case you cannot continue, and yet still want to receive your full compensation, you must stay and wait until the end of the session. You may of course also leave at any time, but you will receive no money for the remaining time of the session. You then have also no more opportunity to obtain further rewards from the apparatus.

Artefact label

The following label was printed and filled out for each tool, core, core fragment and debitage collection.

Participant: ______________________

Session #: _______________________

Date: ___________________________

Core #: __________________________

Not from core? (i.e., HS or AN)

____

Artefact #: _________________________

Success # (if relevant): _________________

Type (circle): Flake-tool *S* *U*

Core-tool

Core fragment

Core

Live coding sheet

The following represents the live coding sheet that was filled out as a paper copy by the experimenter during the sessions and later used as a reference for formulating the video coding of behaviors.

Coding Sheet # ___/___

HIDE FROM PARTICIPANTS AT ALL TIMES

Date	ID	Gender	Session	Core #	Start Time	End Time
__/__/20__		M / F				

	Passive Hammer	Bipolar	Free-handed	Assisted Free-handed	Intent. throw/drop
Attempt=X | Time					

	Friction/rubbing	Pushing with blunt surface	Pushing with sharp edge	Cutting
Attempt=X | Time				

		Technique						Outcome		Use				Core	
		1		2	3	4	5	6	7	8		9		10	
Time	Impossible Flake	Passive hammer		Bipolar	Free-handed	Assisted free-handed	Intent. throw/drop	Flake	Edge on core-tool	Cutting of rope		Other method		Core exhaustion	**Follow-up needed**
		C	E							S	U	S	U		