Expanding the scope of experimental archaeology using the Perception-Process-Product conceptual framework

Cheng Liu*

2023-06-06

Abstract

This paper presents the outline of the Perception-Process-Product ('Triple P') conceptual framework that aims to expand the scope of experimental archaeology. The Triple P framework emphasizes the amplification of multi-level variabilty and the identification of causal relationships of variations across the levels of perception, process, and product. Here we propose the following five basic measures to put the Triple P framework into practice: 1) the acknowledgement of the contribution and limitations of actualistic experiments properly; 2) the normalization the ethological and ethnographic data collection in experimental projects; 3) the involvement of avocational as well as novice participants; 4) the collaboration across labs on a global scale; and 5) the development of an open-access repository for data reuse. ¶

¶ **Keywords:** Experimental archaeology; Ethological analysis; Ethnographical analysis; The curse of knowledge; Collaborative knowledge production

Contents

1 What good is actualistic experiment? 2

2 The ethology and ethnography of stone toolmaking 4

3 The curse of knowledge 6

4 Many places, many voices 7

5 Open science beyond reproducibility 8

6 Acknowledgements 8

References 8

This paper presents the Perception-Process-Product (hereinafter referred to as "Triple P") conceptual framework to expand the scope of experimental archaeology, which tends to center around the reverse engineering of a past technology in a minimal or least-effort manner while ignoring the rich contextual information it affords. Built upon early intellectual principles and practices in

*Department of Anthropology, Emory University, Atlanta, GA, USA; raylc1996@outlook.com

behavioral archaeology (Schiffer, 2010), the Triple P framework aims to **a)** amplify the expression of variability in experimental replicas (product) and their associated behavioral channels (process) as well as sensory experiences (perception) and **b)** better identify the complex interacting relationships across these three levels of variations. To accomplish these two objectives, we advocate the following five measures as integral components of the Triple P framework: 1) acknowledging the contribution and limitations of actualistic experiments properly; 2) normalizing the ethological and ethnographic data collection in experimental projects; 3) encouraging the involvements of avocational as well as novice participants; 4) boosting the collaboration across labs on a global scale; 5) building an open-access repository for data reuse. It is no doubt that strategies of data collection and analysis of a given experimental project should be primarily derived from the research question, which can be legitimately narrow in scope, but the awareness of the rich toolkit available can sometimes inspire researchers to ask questions that are bold and transformative (Schmidt & Marwick, 2020). Here I will mainly leverage the extensive corpus in experimental designs and inferences revolving around stone artifacts to clarify its meaning and demonstrate the necessity and potentials of this framework.

1 What good is actualistic experiment?

The now extensive corpus of experimental archaeology have witnessed the growing principally focused on the generation of knowledge regarding the causal mechanism at behavioral level to explain the variation of material culture (Eren et al., 2016; Lin et al., 2018; Outram, 2008; Reynolds, 1999; Režek et al., 2020). There is no doubt that controlled experiments conducted on stone artifacts (Li et al., 2022), particularly those regarding the fracture mechanics (Cotterell & Kamminga, 1992), provide some foundational and irreplaceable insights on our understanding of the role of lithic technology in prehistory, and unequivocally this line of inquiry should be celebrated and promoted to carry on. Nonetheless, it is oftentimes challenging to directly translate these experimental results into implications of messy human behaviors in the past due to multiple reasons.

Controlled experiments without randomization are not enough to infer causal mechanism, which may be severely biased by factors such as individual differences (Pargeter et al., 2023), allocation concealment (Schulz & Grimes, 2002), and poor recruitment (Fletcher et al., 2012). Randomized Controlled Trials (RCT) *sensu stricto* as practiced in contemporary medicine and behavioral sciences, known for its high cost (e.g., Speich et al., 2019), are extremely rare in experimental archaeology when human participants were involved. Rather, most of our knowledge regarding the past are derived from data sets that can be characterized as Small, Unbalanced, Noisy, but Genuine (SUNG) (Arnaud et al., 2023) produced through expeirments featuring small-sized convenience sample. It has also been a debatable issue whether Randomized Controled Trial (RCT) represents the golden standard of knowledge in both philosophy of science (Cartwright, 2007) and econometrics (Deaton & Cartwright, 2018).

Trade-off between causality and generalizability in experimental design. For example, in a series of recently published experimental studies aiming at understanding the role of language in the transmission of lithic technologies, standardized procedures like video teaching or using bricks as raw materials are very often. One of the major concerns of experimental archaeology design, as in all empirical social sciences, is the validity, namely how good is a particular conclusion or inference approximates the true condition. The concept of validity has multiple dimensions, and one of the most commonly used classification schemes is internal versus external validity. Roe and Just (Roe & Just, 2009: 1266-1267) defined internal validity as "the ability of a researcher to argue that observed correlations are causal" and external validity as "the ability to generalize the relationships found in a study to other persons, times, and settings". This balance between these two validity concepts is an issue that cannot be escaped for all experimental archaeology project designs. In the context of stone tool replication, it can be projected into the debate over the use of machines in knapping, a research design that has received increasing attention in the past decades (Eren et al., 2016). Machine knapping is a typical design with high internal validity but low external validity, which has been proved to provides critical insights into lithic fracture mechanics by identifying potential causal variables at the level of individual stone artifact such as determining which angle of blow or how much force of blow will produce the maximal amount of blade area. All the variables of interest are easy to measure, quantify, and control in a machine knapping setting. Nevertheless, being easy to control is not always a virtue as it essentially eliminates the potential interactions between variables operable in the past and thereby providing misleading results when answering archaeological questions. In addition to the applications of machine knapping, the same problem is also incurred by the introduction of standardized artificial material like bricks or video instruction in teaching experiments. As a rule of thumb, external validity should be given more weight in the design when the research focuses on the behaviors of the users of artifacts while internal validity matters more when it comes to the properties of artifacts themselves.

In the past decades, actualistic experiments becomes more common (Liu & Stout, 2022). Variability as revealed in experiment is crucial and cannot be simply replaced by ethnographic records because the many paleolithic technological components are not displayed in contemporary non-industrial societies, which usually feature technological systems with ground stone artifacts as the target products (Arthur, 2018; Stout, 2002). Statistical techniques for developing causal inference from observational data has also been greatly boosted in recent years (Cunningham, 2021; Hernan & Robins, 2023).

Experimental archaeology is based on the concept of analogy (i.e., the past is at least partially similar to the present in some aspects). It is acknowledged that the validity of this type of analogical inference has long been a subject of debate in archaeology (Chapman & Wylie, 2016; Wylie, 1985), and a comprehensive review of it is beyond the scope of this essay. For instance, presumably in the paleolithic period the development of stone knapping skills mostly happened during childhood and these children grew up in an environment surrounded by habitual stone tool users and makers. In this case, what can knapping teaching experiments involving modern adults who have zero exposure to stone tools inform us about the past learning behavior? It is important to clarify that no experimental project ever intends or claims to provide the perfect reconstruction of the past but rather aims at identifying variables relevant to the question of interest (Stout & Khreisheh, 2015), which can be often ignored through pure deductive reasoning. In the end, all experiments are wrong, but some are useful, and we need more of them.

2 The ethology and ethnography of stone toolmaking

As implied in its name, the implementation of Triple P framework involves the collection of process-level (ethological) and perception-level (ethnographic) data (**Figure** 1), which is critical to address equifinality and multifinality (Hiscock, 2004; Premo, 2010), two daunting challenges in archaeological inference that partially contributed to the discipline-wide paradigm shift in the 1980s (Lake, 2014: 264-265). Equifinality refers to the phenomenon where a similar state or consequence can be achieved through multiple different paths, while multifinality emerges when a similar process can lead to multiple ends. While we cannot fully solve these two problems and accurately reconstruct the past behavioral processes and intentionality simply based on materials remains, context-rich experiments involving the collection of ethological and ethnographic data can help us better document an enlarged range of possible combinations of variation at these three levels and thereby evaluate the probability of certain behavioral mechanisms behind a given archaeological assemblage.

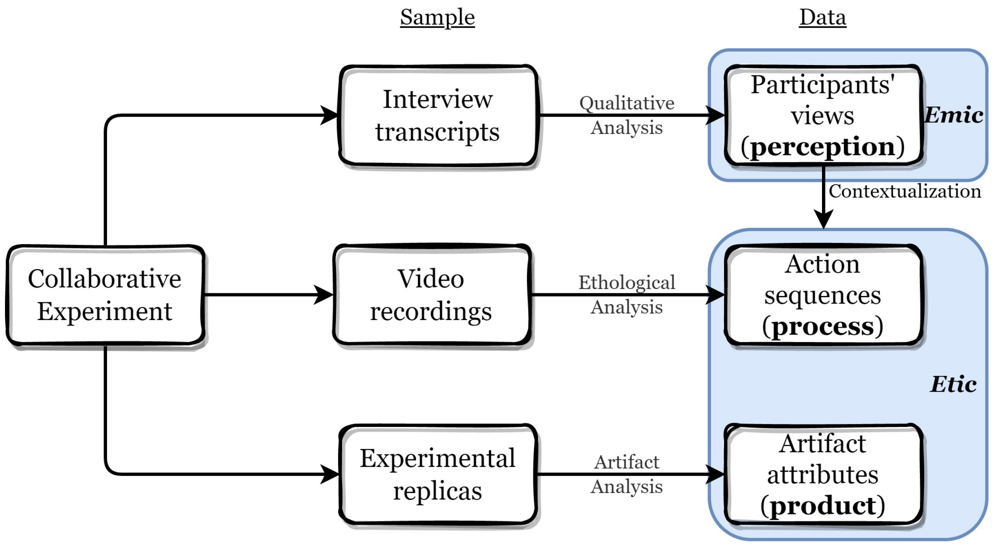

Figure 1: A schematic diagram demonstrating how to operationalize the Perception-Process-Product conceptual framework.

While formal ethological methods that are widely used in the description and analysis of non-human animal behavior (Fragaszy & Mangalam, 2018) still largely fall into oblivion among archaeologists, the attempts of reconstructing behavioral sequences involved in the manufacture of material remains are not infrequent. One such example is cognigram, which was first systematically developed and applied in the archaeological research by Haidle (Miriam N. Haidle, 2010, 2009; Miriam Noël Haidle, 2023; Lombard & Haidle, 2012). Cognigram essentially represents an abstracting process of a series of action sequences achieving a similar goal. This approach is a power and elegant yet limited by the curse of expertise (Hinds, 1999), meaning it cannot handles variability very well. To some extent, it describes the minimal steps to achieve a goal from the perspective of reverse engineering and assume clear causal thinking between each steps, while novices often feature a low planning depth (Opheusden et al., 2023) and a different sets of perception on the causal structure of how certain behaviors will modify the raw materials.

Consequently, we need to accumulate more real-world data by recording a large amount videos of toolmaking and conduct systematical ethogram analysis. With the emergence of new software platforms such as BORIS (Friard & Gamba, 2016), the difficulty of coding has decreased significantly in recent years (**Figure** 2). Here we use action grammar developed by (Stout et al., 2021) as an example. Other coding scheme also exist such as (Mahaney, 2014), or rotation analysis (Muller et al., 2023), or (Cueva-Temprana et al., 2019).

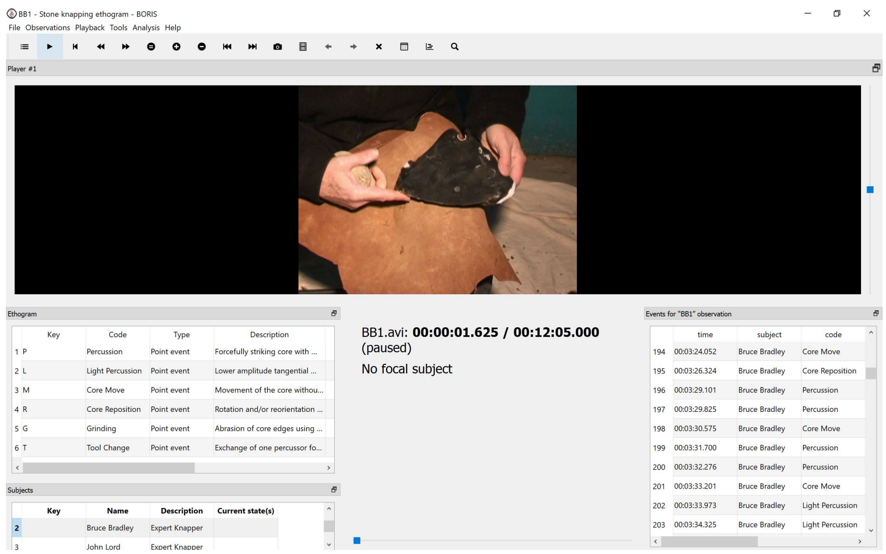

Figure 2: An example of coding Bruce Bradley's handaxe knapping session using the action grammer and BORIS software.

Ethnographies revolving around general archaeological practices (Edgeworth, 2006), experimental archaeology as a field (Reeves Flores, 2012), as well as practices of specific technologies like flintknapping, including both WEIRD (John, C. Whittaker, 2004) and non-WEIRD populations(Arthur, 2018; Stout, 2002), are far from novel. However, it has never been formally recognized as a legitimate research method in experimental archaeology. Echoing with the recent trends of adopting embodied cognition (Varela et al., 2017) in archaeological research (Malafouris, 2013), ethnographic data and methods can reveal hidden information that is otherwise irretrievable and thus should occupy a unique niche in experimental archaeology. This also echoes the post-positivist turn in psychology, a field that is dominated by experimental methods, in the past decades, particularly the emphasis on the value of incorporating qualitative research (Syed & McLean, 2022).

3 The curse of knowledge

We believe that contemporary practices in experimental archaeology, as manifested by the fact the the the majority of scholarly publications are produced as results of experiments conducted by single knapper with a dual identity of researcher, tend to be restrained by the cognitive bias known as the "curse of knowledge" or "curse of expertise". The curse of knowledge refers to the phenomenon that it is extremely challenging for experts to ignore the information that is held by them but not others, particularly novices (Camerer et al., 1989; Hinds, 1999). When the knapping expertise is gradually formed through multiple years of observations and trial-and-error learning, an expert knapper develops some specific ways of strategic planning, motor habits (and their associated impacts on anatomical forms like wrist and elbow), preferences of percussor and raw material types, as well as familiarity of various techniques that become unforgettable. The existence of this cognitive bias is not inherently bad, and these many years of experiences should be appreciated and celebrated by experimental archaeologists. However, what is problematic is that the results of replication experiments conducted by these experienced practitioners, often in settings of single knapper, has been constantly framed as grandiose generalization regarding the evolution of technology and cognition that masks a huge range of technological diversity.

It is more likely for them to come up with ideas that may not be optimal according to the principles of ergonomics. One such example is the the edge angle (Crabtree, 1977)

Experimental archaeology as a scientific method is rooted in the individualistic reverse engineering in the 19th century instead of inter-generation transmission of knapping knowledge that spans several million years (Coles, 1979; Flenniken, 1984; Johnson, 1978; Reeves Flores, 2010; John C. Whittaker, 1994: 54-61).

4 Many places, many voices

Emphasizing variability at its core, the Triple P conceptual framework inherently adopts an collaborative mode of knowledge production, which has been recently advocated in experimental studies (Ranhorn et al., 2020) and museum collection studies (Timbrell, 2022) of stone artifacts.

In addition to the difficulty in coordination and logistics, the facilitation of large-scale collaborations is often hindered by the current system of research evaluation, where usually only the first author and the senior (last/correspondent) author of a peer-reviewed journal paper will be acknowledged as proper contribution.

5 Open science beyond reproducibility

The last step is uploading the data to a open-access repository (Marwick et al., 2017). The building of manufacture can cost (Gilmore et al., 2015; Simon et al., 2015). Following the data sharing principles of FAIR (Wilkinson et al., 2016) and CARE (Carroll et al., 2020)

Given the irreversible nature of archaeological excavations, digitized data, be it text, pictures, or videos, often become the sole evidence that is available for certain research questions. Yet, it has been widely acknowledged that the reuse of archaeological data has not received enough attention among researchers in our discipline (Faniel et al., 2018; Huggett, 2018; Moody et al., 2021). Among many reasons preventing archaeologists from reusing published and digitized data (Sobotkova, 2018), the lack of a standardized practice of and motivation for data sharing is a prominent one (Marwick & Birch, 2018). As stated in the method section, we addressed this issue by sharing the raw data and the code for generating the derived data on an open-access repository. Another major and legitimate concern of archaeological data reuse is their quality. In terms of this aspect, we do acknowledge the limitations of relying on photos when it comes to the more detailed technological analysis of stone artifacts, however, our paper shows that finding the appropriate research questions given the data available is key to revealing new novel insights into the studied topic. Moreover, we believe that this type of research has a strong contemporary relevance due to the continued influence of the COVID-19 on fieldwork-related travel and direct access to archaeological artifacts (Balandier et al., 2022; Ogundiran, 2021).

6 Acknowledgements

This study was supported by a research grant from the Leakey Foundation titled "Inferring skill reproduction from stone artifacts: A middle-range approach" (C. L.).

References

Arnaud, V., Pellegrino, F., Keenan, S., St-Gelais, X., Mathevon, N., Levréro, F., & Coupé, C. (2023). Improving the workflow to crack Small, Unbalanced, Noisy, but Genuine (SUNG) datasets in bioacoustics: The case of bonobo calls. *PLOS Computational Biology*, *19*(4), e1010325. https://doi.org/10.1371/journal.pcbi.1010325

Arthur, K. W. (2018). *The lives of stone tools: Crafting the status, skill, and identity of flintknappers*

(1st edition). University of Arizona Press.

Camerer, C., Loewenstein, G., & Weber, M. (1989). The curse of knowledge in economic settings: An experimental analysis. *Journal of Political Economy*, *97*(5), 1232–1254. https://doi.org/10.1086/261651

Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J., & Hudson, M. (2020). The CARE Principles for Indigenous Data Governance. *Data Science Journal*, *19*(1), 43. https://doi.org/10.5334/dsj-2020-043

Cartwright, N. (2007). Are RCTs the Gold Standard? *BioSocieties*, *2*(1), 11–20. https://doi.org/10.1017/S1745855207005029

Chapman, R., & Wylie, A. (2016). *Evidential reasoning in archaeology*. Bloomsbury Academic.

Coles, J. M. (1979). *Experimental archaeology*. Academic Press.

Cotterell, B., & Kamminga, J. (1992). *Mechanics of pre-industrial technology: An introduction to the mechanics of ancient and traditional material culture*. Cambridge University Press.

Crabtree, D. E. (1977). *The obtuse angle s a functional edge* (D. Ingersoll, J. E. Yellen, & W. MacDonald, Eds.; pp. 38–51). Columbia University Press.

Cueva-Temprana, A., Lombao, D., Morales, J. I., Geribàs, N., & Mosquera, M. (2019). Gestures during knapping: A two-perspective approach to pleistocene technologies. *Lithic Technology*, *44*(2), 74–89. https://doi.org/10.1080/01977261.2019.1587255

Cunningham, S. (2021). *Causal inference: The mixtape*. Yale University Press. https://doi.org/10.2307/j.ctv1c29t27

Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, *210*, 2–21. https://doi.org/10.1016/j.socscimed.2017.12.005

Edgeworth, M. (Ed.). (2006). *Ethnographies of archaeological practice: Cultural encounters, material transformations*. AltaMira Press.

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: Journal of Archaeological, Ethnographic and Experimental Studies*, *8*(2), 103–136. https://doi.org/10.1080/19442890.2016.1213972

Flenniken, J. J. (1984). The past, present, and future of flintknapping: An anthropological perspective. *Annual Review of Anthropology*, *13*(1), 187–203. https://doi.org/10.1146/annurev.an.13.100184.001155

Fletcher, B., Gheorghe, A., Moore, D., Wilson, S., & Damery, S. (2012). Improving the recruitment activity of clinicians in randomised controlled trials: a systematic review. *BMJ Open*, *2*(1), e000496. https://doi.org/10.1136/bmjopen-2011-000496

Fragaszy, D. M., & Mangalam, M. (2018). *Chapter Five - Tooling* (M. Naguib, L. Barrett, S. D. Healy, J. Podos, L. W. Simmons, & M. Zuk, Eds.; Vol. 50, pp. 177–241). Academic Press. https://doi.org/10.1016/bs.asb.2018.01.001

Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, *7*(11), 1325–1330. https://doi.org/10.1111/2041-210X.12584

Gilmore, R., Adolph, K., Millman, D., Steiger, L., & Simon, D. (2015). Sharing displays and data from vision science research with databrary. *Journal of Vision*, *15*(12), 280. https://doi.org/10.1167/15.12.280

Haidle, Miriam 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. https://doi.org/10.1086/650295

Haidle, Miriam N. (2009). How to think a simple spear. In S. A. de Beaune, F. L. Coolidge, & T. Wynn (Eds.), *Cognitive archaeology and human evolution* (pp. 57–73). Cambridge University Press.

Haidle, Miriam Noël. (2023). *Cognigrams: Systematically reconstructing behavioral architectures as a basis for cognitive archaeology* (T. Wynn, K. A. Overmann, & F. L. Coolidge, Eds.; p. C12S1C12S8). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780192895950.013.12

Hernan, M. A., & Robins, J. M. (2023). *Causal inference: What if*. CRC Press.

Hinds, P. J. (1999). The curse of expertise: The effects of expertise and debiasing methods on prediction of novice performance. *Journal of Experimental Psychology: Applied*, *5*, 205–221. https://doi.org/10.1037/1076-898X.5.2.205

Hiscock, P. (2004). Slippery and Billy: Intention, Selection and Equifinality in Lithic Artefacts. *Cambridge Archaeological Journal*, *14*(1), 71–77. https://doi.org/10.1017/S0959774304230050

Johnson, L. L. (1978). A history of flint-knapping experimentation, 1838-1976 [and comments and reply]. *Current Anthropology*, *19*(2), 337–372. https://doi.org/10.1086/202078

Lake, M. W. (2014). Trends in Archaeological Simulation. *Journal of Archaeological Method and Theory*, *21*(2), 258–287. https://doi.org/10.1007/s10816-013-9188-1

Li, L., Lin, S. C., McPherron, S. P., Abdolahzadeh, A., Chan, A., Dogandžić, T., Iovita, R., Leader, G. M., Magnani, M., Rezek, Z., & 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*. https://doi.org/10.1007/s10816-022-09586-2

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. https://doi.org/10.1007/s10816-017-9351-1

Liu, C., & Stout, D. (2022). Inferring cultural reproduction from lithic data: A critical review. *Evolutionary anthropology*. https://doi.org/10.1002/evan.21964

Lombard, M., & Haidle, M. N. (2012). Thinking a Bow-and-arrow Set: Cognitive Implications of Middle Stone Age Bow and Stone-tipped Arrow Technology. *Cambridge Archaeological Journal*, *22*(2), 237–264. https://doi.org/10.1017/S095977431200025X

Mahaney, R. A. (2014). Exploring the complexity and structure of acheulean stoneknapping in relation to natural language. *PaleoAnthropology*, *2014*, 586606. https://doi.org/10.4207/PA.2014.ART90

Malafouris, L. (2013). *How things shape the mind: A theory of material engagement*. The MIT Press.

Marwick, B., d'Alpoim Guedes, J., Barton, C. M., Bates, L. A., Baxter, M., Bevan, A., Bollwerk, E. A., Bocinsky, R. K., Brughmans, T., Carter, A. K., Conrad, C., Contreras, D. A., Costa, S., Crema, E. R., Daggett, A., Davies, B., Drake, B. L., Dye, T. S., France, P., … Wren, C. D. (2017). Open science in archaeology. *SAA Archaeological Record*, *17*(4), 8–14. http://onlinedigeditions.com/publication/?i=440506

Muller, A., Shipton, C., & Clarkson, C. (2023). The Proceduralization of Hominin Knapping Skill: Memorizing Different Lithic Technologies. *Cambridge Archaeological Journal*, 1–18. https://doi.org/10.1017/S0959774323000070

Opheusden, B. van, Kuperwajs, I., Galbiati, G., Bnaya, Z., Li, Y., & Ma, W. J. (2023). Expertise increases planning depth in human gameplay. *Nature*, 1–6. https://doi.org/10.1038/s41586-023-06124-2

Outram, A. K. (2008). Introduction to experimental archaeology. *World Archaeology*, *40*(1), 1–6. https://www.jstor.org/stable/40025310

Pargeter, J., Liu, C., Kilgore, M. B., Majoe, A., & Stout, D. (2023). Testing the Effect of Learning Conditions and Individual Motor/Cognitive Differences on Knapping Skill Acquisition. *Journal* *of Archaeological Method and Theory*, *30*(1), 127–171. https://doi.org/10.1007/s10816-022-09592-4

Premo, L. S. (2010). Equifinality and explanation: Thoughts on the role of agent-based modeling in postpositivist archaeology. In A. Costopoulos & M. W. Lake (Eds.), *Simulating Change: Archaeology Into the Twenty-first Century* (pp. 28–37). University of Utah Press.

Ranhorn, K. L., Pargeter, J., & Premo, L. S. (2020). Investigating the evolution of human social learning through collaborative experimental archaeology. *Evolutionary Anthropology: Issues, News, and Reviews*, *29*(2), 53–55. https://doi.org/10.1002/evan.21823

Reeves Flores, J. (2012). *Experimental archaeology: an ethnography of its perceived value and impact in archaeological research* [PhD thesis]. https://ore.exeter.ac.uk/repository/handle/10871/9041

Reeves Flores, J. (2010). *Creating a history of experimental archaeology* (D. Millson, Ed.; pp. 29–45). Oxbow Books.

Reynolds, P. J. (1999). *The nature of experiment in archaeology* (A. Harding, Ed.; p. 156162). Oxbow Books.

Režek, Ž., Holdaway, S. J., Olszewski, D. I., Lin, S. C., Douglass, M., McPherron, S. P., Iovita, R., Braun, D. R., & Sandgathe, D. (2020). Aggregates, Formational Emergence, and the Focus on Practice in Stone Artifact Archaeology. *Journal of Archaeological Method and Theory*, *27*(4), 887–928. https://doi.org/10.1007/s10816-020-09445-y

Roe, B. E., & Just, D. R. (2009). Internal and external validity in economics research: Tradeoffs between experiments, field experiments, natural experiments, and field data. *American Journal of Agricultural Economics*, *91*(5), 1266–1271. https://www.jstor.org/stable/20616293

Schiffer, M. B. (2010). *Behavioral Archaeology: Principles and Practice*. Routledge.

Schmidt, S. C., & Marwick, B. (2020). Tool-Driven Revolutions in Archaeological Science. *Journal of Computer Applications in Archaeology*, *3*(1), 1832. https://doi.org/10.5334/jcaa.29

Schulz, K. F., & Grimes, D. A. (2002). Allocation concealment in randomised trials: defending against deciphering. *The Lancet*, *359*(9306), 614–618. https://doi.org/10.1016/S0140-6736(02)07750-4

Simon, D. A., Gordon, A. S., Steiger, L., & Gilmore, R. O. (2015). *Databrary: Enabling sharing and reuse of research video*. 279280. https://doi.org/10.1145/2756406.2756951

Speich, B., Schur, N., Gryaznov, D., Niederhäusern, B. von, Hemkens, L. G., Schandelmaier, S., Amstutz, A., Kasenda, B., Pauli-Magnus, C., Ojeda-Ruiz, E., Tomonaga, Y., McCord, K., Nordmann, A., Elm, E. von, Briel, M., Schwenkglenks, M., & Groups, a. collaboration of the M. (MAking. R. T. A. and A. (Adherence. to S. P. I. Re. for interventional trials). S. (2019). Resource use, costs, and approval times for planning and preparing a randomized clinical trial before and after the implementation of the new Swiss human research legislation. *PLOS ONE*, *14*(1), e0210669. https://doi.org/10.1371/journal.pone.0210669

Stout, D. (2002). Skill and cognition in stone tool production: An ethnographic case study from irian jaya. *Current Anthropology*, *43*(5), 693–722. https://doi.org/10.1086/342638

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). https://doi.org/10.1038/s41598-021-92992-5

Stout, D., & Khreisheh, N. (2015). Skill Learning and Human Brain Evolution: An Experimental Approach. *Cambridge Archaeological Journal*, *25*(4), 867–875. https://doi.org/10.1017/S0959774315000359

Syed, M., & McLean, K. C. (2022). Disentangling paradigm and method can help bring qualitative research to post-positivist psychology and address the generalizability crisis. *Behavioral and Brain Sciences*, *45*, e32. https://doi.org/10.1017/S0140525X21000431

Timbrell, L. (2022). A collaborative model for lithic shape digitization in museum settings. *Lithic Technology*, *0*(0), 1–12. https://doi.org/10.1080/01977261.2022.2092299

Varela, F. J., Thompson, E., & Rosch, E. (2017). *The Embodied Mind: Cognitive Science and Human Experience* (revised edition). The MIT Press.

Whittaker, John C. (1994). *Flintknapping: Making and Understanding Stone Tools*. University of Texas Press.

Whittaker, John, C. (2004). *American Flintknappers: Stone Age Art in the Age of Computers*. University of Texas Press.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Silva Santos, L. B. da, Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 160018. https://doi.org/10.1038/sdata.2016.18

Wylie, A. (1985). The reaction against analogy. *Advances in Archaeological Method and Theory*, *8*, 63–111. https://doi.org/10.1016/B978-0-12-003108-5.50008-7