ANTIQUITY
a review of world archaeology

CAMBRIDGE
UNIVERSITY PRESS

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

Abstract:	This paper presents the Perception-Process-Product ('Triple P') framework that aims to expand the scope of experimental archaeology. The Triple P framework emphasizes multi-level variation and interactions across the levels of perception, process, and product to provide a more grounded and richer explanation of the past archaeological record. It consists of three principles: 1) acknowledging the inherent trade-off between control and generalizability in naturalistic research design; 2) encouraging collaborative projects that involve geographically diverse and non-traditional research participants such as hobbyists and novices; 3) adopting a workflow that normalizes the collection and curation of ethological and ethnographic data in experimental projects.

SCHOLARONE™
Manuscripts

Abstract: This paper presents the Perception-Process-Product ('Triple P') framework that aims to expand the scope of experimental archaeology. The Triple P framework emphasizes multi-level variation and interactions across the levels of perception, process, and product to provide a more grounded and richer explanation of the past archaeological record. It consists of three principles: 1) acknowledging the inherent trade-off between control and generalizability in naturalistic research design; 2) encouraging collaborative projects that involve geographically diverse and non-traditional research participants such as hobbyists and novices; 3) adopting a workflow that normalizes the collection and curation of ethological and ethnographic data in experimental projects.

Introduction

This paper presents the Perception-Process-Product (hereafter 'Triple P') conceptual framework to expand the scope of experimental archaeology. The field has long tended to adopt the principle of Occam's razor (e.g., Domínguez-Rodrigo, 2008; Reeves et al., 2009), whether explicitly or implicitly, and center around the reverse engineering of a past technology in a minimal or least-effort manner while ignoring the rich contextual information it affords. Nevertheless, Occam's razor can be insufficient to infer the preferences of 'irrational' agents possessing incomplete information (Mindermann & Armstrong, 2018) in tool design and use, and the two conditions described here provide a better approximation of past humans displaying extensive cultural variation as opposed to the assumption of omniscient *Homo economicus* for most anthropologists (Henrich et al., 2001). In the evolution of technology, it is rather common that opaque causal perception and its resulting tendency of over-imitation (e.g., Sylvia's Recipe, Gergely & Csibra, 2006) together with path dependence (e.g., QUERTY keyboard design, Kafaee et al., 2022) can lead to the widespread and long-lasting reproduction of technological solutions that are neither optimal in functional efficiency nor minimal in manufacture complexity.

Built upon the *Homo economicus* critique and early works in behavioural archaeology (Schiffer, 2010), here I propose the Triple P framework, which aims to **a)** amplify the expression of variation in experimental replicas (product) and their associated behavioural channels (process) as well as sensory experiences (perception) by experiments in diverse

30　contexts and **b)** better identify the complex interacting relationships across these three

31　levels of variations in real-world conditions. To accomplish these two objectives, I advocate

32　the following three principles as integral components of the Triple P framework, which

33　requires **1)** acknowledging the inherent trade-off between control and generalizability in

34　naturalistic research design and **2)** encouraging collaborative projects that involve

35　geographically diverse and non-traditional research participants such as hobbyists and

36　novices. These two principles are developed to advocate a pluralistic approach to the

37　explanation of complex variation, which has received more attention from evolutionary

38　anthropology (Antón & Kuzawa, 2017) to cognitive science (Barrett, 2020), instead of

39　treating the optimization-based research agenda as a panacea. The second principle

40　particularly allows researchers to develop research questions that are also meaningful to

41　descendant communities through respectful conversation and collaboration (Montgomery

42　& Fryer, 2023). The Triple P framework also **3)** adopts a workflow that normalizes the

43　collection and curation of ethological and ethnographic data in experimental projects. It is

44　acknowledged that strategies of data collection and analysis of a given experimental project

45　should be primarily derived from the research question, but the awareness of the rich

46　toolkit available can sometimes inspire researchers to ask questions that are bold and

47　transformative (Schmidt & Marwick, 2020). Here I will leverage the extensive corpus in

48　experimental designs and inferences revolving around stone artefacts to clarify its meaning

49　and demonstrate the necessity and potential of this framework.

What good is naturalistic experimentation?

51　The trade-off between causal inference (aka 'internal validity') and generalization (aka

52　'external validity') forms a central issue in experimental design across different disciplines

53　(Eren et al., 2016; Roe & Just, 2009: 1266-1267). Even in fields known for their

54　development of rigorous and well-controlled experimental methods such as cognitive

55　psychology and neuroscience, researchers have started to use naturalistic stimuli more

56　frequently and advocate a paradigm shift to semi-controlled experiments due to the

57　generalizability crisis, namely the prevailing mismatch between phenomenon of interest

58　and measured variables in psychological science (Nastase et al., 2020; Shamay-Tsoory &

59   Mendelsohn, 2019; Sonkusare et al., 2019; Yarkoni, 2022). In contrast, the past decades

60   have witnessed experimental archaeology's growing research interests focusing on the

61   causal mechanism at the behavioural level in the explanation of material culture variation

62   (Eren et al., 2016; Lin et al., 2018). In the context of stone artefact replication, one typical

63   research design emphasizing causality over generalizability is the use of knapping

64   machines/robots (Li et al., 2022; Pfleging et al., 2019), which has helped map out the

65   physical constraints of stone artefact manufacture and use through the identification of

66   causal relationships between input (force, exterior platform angle, platform depth, etc.) and

67   outcome variables (flake size, flake shape, wear formation, etc.). All variables of interest in

68   this setting are relatively easy to measure, quantify, and control, but this type of design can

69   be insufficient in inferring how context-generic principles interact in a particular context as

70   reflected in real-world conditions. Similarly, standardized artificial materials like bricks

71   (Lombao et al., 2017) or foam blocks (Schillinger et al., 2016) have been used to

72   standardize materials and/or reduce learning demands in experimental studies focusing on

73   the transmission of lithic technologies, with implications for the generalizability of results

74   (Liu et al., 2023). In real-world knapping, each rock has a different shape and often

75   different physical properties such as inner cracks and inclusions, and this heterogeneity

76   itself represents a critical variable in cultural transmission and skill development (Proffitt

77   et al., 2022).

78   On the other hand, naturalistic experiments pay more attention to how experimental

79   insights can be generalized to archaeological samples by incorporating authentic materials

80   and plausible social settings with a certain degree of compromised control (Outram, 2008).

81   Back to the cases of cultural transmission experiments, a naturalistic experiment would

82   involve the use of natural rocks with varied morphology instead of standardized artificial

83   materials as well as human demonstrators instead of videos of knapping instruction,

84   despite the fact that the latter will remain consistent across individuals. Unlike controlled

85   experiments, variation could be easily observed in naturalistic experiments by design. This

86   feature is crucial and cannot be simply replaced by ethnographic records, because many

87   palaeolithic technological components do not have analogues in contemporary non-

88   industrial societies (e.g., Arthur, 2018; Stout, 2002). While uncontrolled variation has

89 traditionally been viewed as highly problematic, statistical techniques for developing

90 causal inference from observational data, of the kind produced by naturalistic experiments,

91 have also been greatly boosted in epidemiology and economics in recent years

92 (Cunningham, 2021; Hernan & Robins, 2023). Naturalistic experiment can serve a heuristic

93 role in hypothesis generation, aligning with the perspective of Lin et al. (2018: 680-681),

94 who proposed that the interaction between naturalistic and controlled experiment

95 "operates in a cyclical form of induction and deduction."

Many places, many voices

97 Contemporary practices in experimental archaeology, as manifested by the fact that a

98 majority of scholarly publications are produced as results of experiments conducted by a

99 single knapper with a dual identity of researcher (Whittaker, 2004), tend to be restrained

100 by the cognitive bias known as the 'curse of knowledge' or 'curse of expertise'. This

101 psychological term originally refers to the phenomenon that it is extremely challenging for

102 experts to ignore the information that is held by them but not others, particularly novices

103 (Hinds, 1999), but it has further implications for the sample representativeness in

104 experimental archaeology. When the knapping expertise is gradually formed through

105 multiple years of observations and trial-and-error learning, an expert knapper develops

106 some specific ways of strategic planning, motor habits (and their associated impacts on

107 anatomical forms like wrist and elbow), preferences of percussor and raw material types,

108 as well as familiarity of various techniques that become unforgettable (Moore, 2020: 654).

109 The existence of this cognitive bias is not inherently bad, and these many years of

110 experience should be appreciated and celebrated by experimental archaeologists.

111 However, what is problematic is that the results of replication experiments conducted by

112 these experienced practitioners, often in settings of single knapper, has been constantly

113 framed as generalization regarding the evolution of technology and cognition that masks a

114 vast range of technological diversity.

115 Modern flintknapping techniques, as a research subject and a scientific method, originated

116 from hobbyists' individualistic trials of reverse engineering during the 19th century (Coles,

1979; Flenniken, 1984; Johnson, 1978; Whittaker, 1994: 54-61). Hobbyist knappers represent a huge repertoire of technological knowledge that does not fully overlap with what is acquired by academic knappers. They tend to generate ideas that may appear to be counter-intuitive at first glance for academics. One such example is the utility of obtuse edge angle as demonstrated by Don Crabtree (1977), a mostly self-educated flintknapper yet one of the most important figures in experimental archaeology. In his experiment, Crabtree demonstrated the excellent performance of blade dorsal ridge on tasks like shaving and cutting hard materials, challenging the traditional perspective on producing sharp lateral edges as the sole purpose of stone toolmaking and shedding light on future functional reconstruction through the use-wear analysis. It is rather unfortunate that collaborations between academics and hobbyists are less common than expected due to their complicated and uneasy relationships as detailed in Whittaker's (2004) ethnography. Likewise, novices' lack of expertise also helps to mitigate the 'curse of knowledge' bias that may hinder expert knappers. Their involvement can potentially lead to the discovery of alternative methods, techniques, and interpretations that may have been overlooked by experts.

Emphasizing variation at its core, the Triple P conceptual framework recognizes that experimental archaeology can greatly benefit from diverse perspectives (Pargeter et al., 2023: 164) and thereby inherently adopts a collaborative mode of knowledge production, which has been recently advocated in experimental studies (Liu & Stout, 2023; Ranhorn et al., 2020) and museum collection studies (Timbrell, 2023) of stone artefacts. Furthermore, the Triple P framework acknowledges that communities living in specific geographical areas possess unique insights and understanding of their cultural heritage. This emphasis on team efforts and inclusivity allows for a more complete understanding of the non-utilitarian or unexpected aspects of raw material selection (Arthur, 2021), pre-treatment (Maloney & Street, 2020), production (Griffin et al., 2013), and use (Martellotta et al., 2022) across different regions. Through ethical collaborations with those knapping practitioners in non-industrial societies in the research process, the framework allows their voices to be heard and their contributions to be acknowledged. This not only enhances the quality of research outcomes but also fosters a sense of ownership and pride within these

147 communities, strengthening the connection between archaeological research and the

148 people it directly affects (Montgomery & Fryer, 2023).

The Triple P framework in action

150 The implementation of the Triple P framework involves the collection of process-level

151 (ethological) and perception-level (ethnographic) data (**Fig. 1**), which is critical to address

152 equifinality and multifinality (Nami, 2010; Premo, 2010), two daunting challenges in

153 archaeological inference. Equifinality refers to situations in which a similar state or

154 consequence can be achieved through multiple different paths, while multifinality emerges

155 when a similar process can lead to multiple ends. While we cannot fully solve these two

156 problems and accurately reconstruct the past behavioural processes simply based on

157 materials remains, context-rich experiments involving the collection of ethological and

158 ethnographic data can help us better document an enlarged range of possible combinations

159 of variation and draw a more informed inference (Reynolds, 1999). The importance of

160 specifying and documenting the context information of both the experiment and the

161 phenomenon of interest has also been recently highlighted in psychological sciences
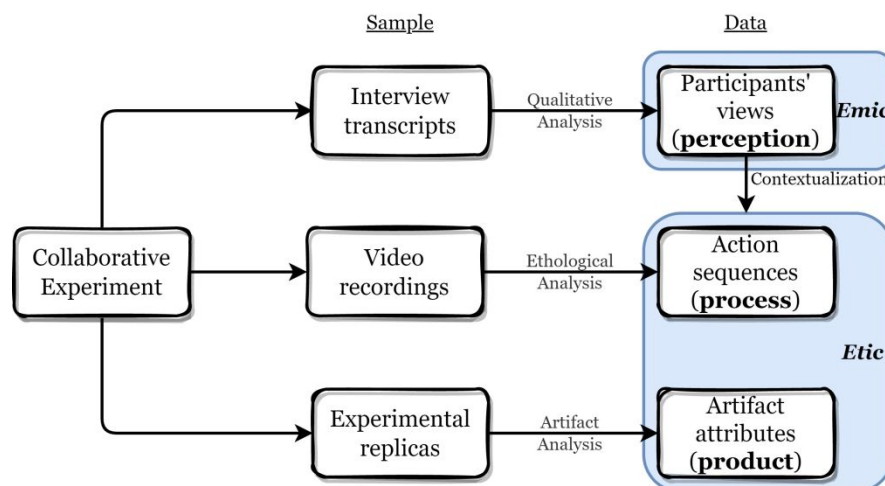
162 (Holleman et al., 2020).

163

164 *Fig.1. A schematic diagram demonstrating how to operationalize the Perception-Process-*

165 *Product conceptual framework.*

Product-level data

Traditionally speaking, the product-level data, namely the documentation and analysis of replicas, form the sole research subject of experimental archaeology and serve as the tangible foundation for analogical inference in the interpretation of archaeological materials. It can exist in the form of spreadsheets containing detailed technological attributes, photos and illustrations, or high-resolution 3D scans of individual artefacts or a whole assemblage. No particular modification regarding the collection procedure of product-level data is required in the context of the Triple P framework, although the definition of variables measured and the documentation techniques (models of camera/scanners, light setting, processing software version, etc.) should be always available in the relevant meta-data. I also strongly recommend adopting good habits in spreadsheet data organization (Broman & Woo, 2018).

Process-level data

While systematic behavioural coding methods widely used in the study of non-human animal behaviour (Fragaszy & Mangalam, 2018) are still largely neglected among archaeologists, attempts to reconstruct behavioural sequences involved in the manufacture of material remains are not infrequent. One such example is the cognigram, which was first systematically developed and applied in archaeological research by Haidle (Haidle, 2009, 2014). A cognigram is a graphical representation of the reconstructed behaviour behind archaeological artefacts in chronological order of appearance (Haidle, 2014), which essentially represents an abstracting process of a series of action sequences achieving a similar goal. This approach provides an elegant descriptive methodology yet is limited by its normative and analytical orientation, meaning it cannot handle variation very well. To some extent, it describes the minimal steps to achieve a goal from the perspective of reverse engineering and reflects the analyst's own causal perception. However, this may be biased because 1) certain causal insights in stone fracture mechanics remain opaque to academic knappers until they are revealed through controlled experiments by Dibble and his colleagues (Li et al., 2022) 2) ethnographic studies demonstrated that expert non-

194   academic practitioners can have a different set of causal understanding (Harris et al.,

195   2021).

196   Consequently, we need to accumulate more real-world data by recording a large number of

197   toolmaking videos and conducting systematic ethogram analysis. With the emergence of

198   new software platforms such as BORIS (Friard & Gamba, 2016), the difficulty of coding has

199   decreased significantly in recent years (**Fig. 2**). Here I use a modified version of action

200   grammar developed by (Stout et al., 2021) as an example, among multiple coding schemes

201   featuring different research focus (Muller et al., 2023) or granularity (Cueva-Temprana et

202   al., 2019; Mahaney, 2014; Roux & David, 2005). The knapping action recorded in videos can

203   be coded following the ethogram presented in **Table 1**. Depending on the original research

204   question, sequences of coded actions can then be used in further analysis, such as

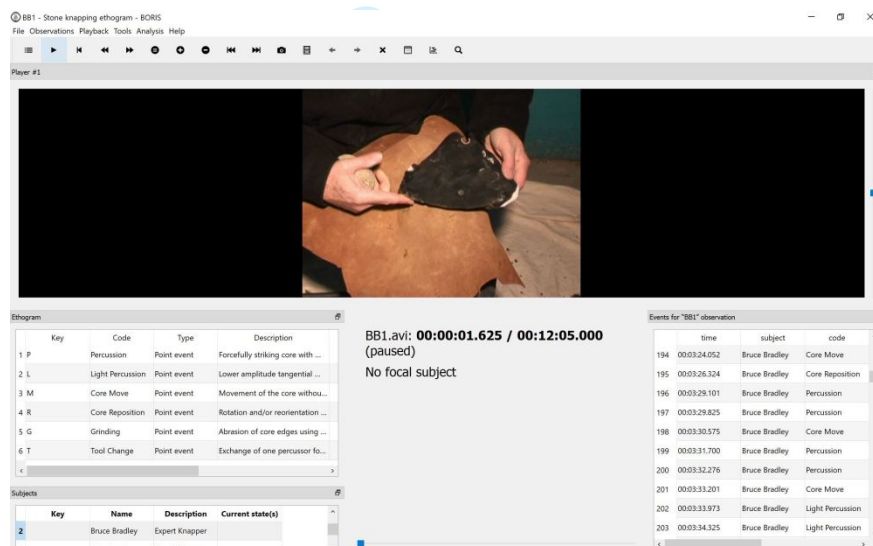205   complexity (Stout et al., 2021), similarity (Mobbs et al., 2021), etc.

206

207   *Fig.2. An example of coding a handaxe knapping session using the BORIS software.*

208

209

210

211 *Table 1. A modified version of the original action grammar presented in (Stout et al., 2021)*

Action	Definition
Percussion	Forcefully striking core with percussor (hammerstone or antler billet) in such a way as to potentially remove a flake
Light Percussion	Lower amplitude tangential strike to the tool edge of the kind often employed for platform preparation
Core Move	Movement of the core without a change in grip.
Core Reposition	Rotation and/or reorientation of the core involving repositioning of the hand.
Grinding	Abrasion of core edges using a hammerstone
Tool Change	Exchange of one percussor for another

212 **Perception-level data**

213 Ethnographies revolving around experimental archaeology as a field (Reeves Flores, 2012),
214 as well as practices of specific technologies like flintknapping, including contemporary U.S.
215 hobbyists (Whittaker, 2004) and knapping practitioners in various non-industrial societies
216 (Arthur, 2018; Stout, 2002), are far from novel. However, ethnography has never been
217 formally recognized as a legitimate research method in experimental archaeology. Echoing
218 with the recent trends of adopting embodied cognition (Varela et al., 2017) in
219 archaeological research (Malafouris, 2013), ethnographic data and methods can reveal
220 hidden information (e.g., intention, phenomenology) that is otherwise irretrievable and
221 thus should occupy a unique niche in experimental archaeology. This also echoes the post-
222 positivist turn in psychology in the past decades, particularly the emphasis on the value of
223 incorporating qualitative research (Stout, 2021; Syed & McLean, 2022; Weger et al., 2019).

224 Through participant observation, interviews, and detailed field notes, ethnography can
225 capture the subtle nuances of perception, such as sensory experiences, social interactions,
226 and cultural meanings associated with the experimental activities (Gowlland, 2019).
227 Compared with the ethological methods, the interview questions and participant
228 observation in ethnographic methods feature an even higher degree of freedom and rely

229   more heavily on the research question as well as ad-hoc interaction. One potential
230   application of ethnographic methods in experimental archaeology of stone artefacts is
231   asking knappers about the intentions of each action and see how it matches with the
232   results as revealed by lithic analysis of replicas, which can provide crucial contextual
233   information addressing the issues of equifinality and multifinality in the formation of lithic
234   assemblage.

Multi-level data curation

236   The comparative study and large-scale synthesis of variation data require the building of
237   centralized, open-access, and carefully curated data infrastructure, which unfortunately
238   still does not exist yet in experimental archaeology. Among the three dimensions of the
239   Triple P framework, the product-level data are usually stored in the format of
240   spreadsheets, photos, and 3D models, and the perception-level data formats mainly include
241   audio files and their transcribed texts, whereas videos are the main vector of process-level
242   data, a rather non-traditional data format in archaeological research featuring the highest
243   file size compared with the other two. As such, following data sharing principles of FAIR
244   (Wilkinson et al., 2016) and CARE (Carroll et al., 2020), the Triple P framework
245   recommends Databrary (Simon et al., 2015), a web-based library originally designed for
246   developmental scientists, as the main data curation platform, where researchers can freely
247   upload video files and related metadata that can connect with different types of data within
248   the same project.

Conclusion

250   Through the broadening of traditional data types and recording methods revolving around
251   experimental replicas *per se*, the Triple P conceptual framework allows the amplified
252   multiscale expression of material cultural variation. It is also compatible with many
253   theoretical orientations, ranging from behavioural archaeology (emphasis on video
254   recording of behavioural processes) through evolutionary archaeology (emphasis on the
255   amplification of variation) to post-processual archaeology (emphasis on perception

through ethnography). In terms of its research practice, it embraces a collaborative mode of knowledge production by involving a more diverse pool of stakeholders. The innovativeness, flexibility, and inclusiveness of the Triple P conceptual framework has enormous potential in redefining what can be and what should be studied by experimental archaeology as a field and thereby contributing to a better understanding of our deep past.

Acknowledgements

References

Antón, S. C., & Kuzawa, C. W. (2017). Early homo, plasticity and the extended evolutionary synthesis. *Interface Focus*, *7*(5), 20170004. https://doi.org/10.1098/rsfs.2017.0004

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

Arthur, K. W. (2021). Material Scientists: Learning the Importance of Colour and Brightness from Lithic Practitioners. *Cambridge Archaeological Journal*, *31*(2), 293–304. https://doi.org/10.1017/S0959774320000347

Barrett, H. C. (2020). Towards a Cognitive Science of the Human: Cross-Cultural Approaches and Their Urgency. *Trends in Cognitive Sciences*, *24*(8), 620–638. https://doi.org/10.1016/j.tics.2020.05.007

Broman, K. W., & Woo, K. H. (2018). Data organization in spreadsheets. *The American Statistician*, *72*(1), 2–10. https://doi.org/10.1080/00031305.2017.1375989

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.,

281 Anderson, J., & Hudson, M. (2020). The CARE Principles for Indigenous Data Governance.

282 *Data Science Journal*, *19*(1), 43. https://doi.org/10.5334/dsj-2020-043

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

284 Crabtree, D. E. (1977). *The obtuse angle as a functional edge* (D. Ingersoll, J. E. Yellen, & W.

285 MacDonald, Eds.; pp. 38–51). Columbia University Press.

286 Cueva-Temprana, A., Lombao, D., Morales, J. I., Geribàs, N., & Mosquera, M. (2019). Gestures

287 during knapping: A two-perspective approach to pleistocene technologies. *Lithic*

288 *Technology*, *44*(2), 74–89. https://doi.org/10.1080/01977261.2019.1587255

289 Cunningham, S. (2021). *Causal inference: The mixtape*. Yale University Press.

290 https://doi.org/10.2307/j.ctv1c29t27

291 Domínguez-Rodrigo, M. (2008). Conceptual premises in experimental design and their

292 bearing on the use of analogy: An example from experiments on cut marks. *World*

293 *Archaeology*, *40*(1), 67–82. https://doi.org/10.1080/00438240701843629

294 Eren, M. I., Lycett, S. J., Patten, R. J., Buchanan, B., Pargeter, J., & O'Brien, M. J. (2016). Test,

295 model, and method validation: The role of experimental stone artifact replication in

296 hypothesis-driven archaeology. *Ethnoarchaeology: Journal of Archaeological, Ethnographic*

297 *and Experimental Studies*, *8*(2), 103–136.

298 https://doi.org/10.1080/19442890.2016.1213972

299 Flenniken, J. J. (1984). The past, present, and future of flintknapping: An anthropological

300 perspective. *Annual Review of Anthropology*, *13*(1), 187–203.

301 https://doi.org/10.1146/annurev.an.13.100184.001155

302 Fragaszy, D. M., & Mangalam, M. (2018). *Chapter Five - Tooling* (M. Naguib, L. Barrett, S. D.

303 Healy, J. Podos, L. W. Simmons, & M. Zuk, Eds.; Vol. 50, pp. 177–241). Academic Press.

304 https://doi.org/10.1016/bs.asb.2018.01.001

305   Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software

306   for video/audio coding and live observations. *Methods in Ecology and Evolution*, *7*(11),

307   1325–1330. https://doi.org/10.1111/2041-210X.12584

308   Gergely, G., & Csibra, G. (2006). *Sylvia's recipe: The role of imitation and pedagogy in the*

309   *transmission of cultural knowledge* (S. C. Levinson & N. J. Enfield, Eds.; pp. 229–255). Berg

310   Publishers.

311   Gowlland, G. (2019). The sociality of enskilment. *Ethnos*, *84*(3), 508–524.

312   https://doi.org/10.1080/00141844.2018.1455726

313   Griffin, D., Freedman, D. L., Nicholson, B., McConachie, F., & Parmington, A. (2013). The

314   koorong project: Experimental archaeology and wurundjeri continuation of cultural

315   practices. *Excavations, Surveys and Heritage Management in Victoria*, *2*, 5965.

316   Haidle, M. N. (2009). How to think a simple spear. In S. A. de Beaune, F. L. Coolidge, & T.

317   Wynn (Eds.), *Cognitive archaeology and human evolution* (pp. 57–73). Cambridge

318   University Press.

319   Haidle, M. N. (2014). Building a bridgean archeologist's perspective on the evolution of

320   causal   cognition.   *Frontiers   in   Psychology*,   *5*.

321   https://www.frontiersin.org/articles/10.3389/fpsyg.2014.01472

322   Harris, J. A., Boyd, R., & Wood, B. M. (2021). The role of causal knowledge in the evolution of

323   traditional   technology.   *Current   Biology*,   *31*(8),   1798–1803.e3.

324   https://doi.org/10.1016/j.cub.2021.01.096

325   Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In

326   Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. *American*

327   *Economic Review*, *91*(2), 73–78. https://doi.org/10.1257/aer.91.2.73

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

329   Hinds, P. J. (1999). The curse of expertise: The effects of expertise and debiasing methods

330   on prediction of novice performance. *Journal of Experimental Psychology: Applied*, *5*, 205–

331   221. https://doi.org/10.1037/1076-898X.5.2.205

332   Holleman, G. A., Hooge, I. T., Kemner, C., & Hessels, R. S. (2020). The 'real-world

333   approach'and its problems: A critique of the term ecological validity. *Frontiers in*

334   *Psychology*, *11*, 721.

335   Johnson, L. L. (1978). A history of flint-knapping experimentation, 1838-1976 [and

336   comments   and   reply].   *Current   Anthropology*,   *19*(2),   337–372.

337   https://doi.org/10.1086/202078

338   Kafaee, M., Daviran, E., & Taqavi, M. (2022). The QWERTY keyboard from the perspective of

339   the Collingridge dilemma: lessons for co-construction of human-technology. *AI & SOCIETY*.

340   https://doi.org/10.1007/s00146-022-01573-1

341   Li, L., Lin, S. C., McPherron, S. P., Abdolahzadeh, A., Chan, A., Dogandžić, T., Iovita, R., Leader,

342   G. M., Magnani, M., Rezek, Z., & Dibble, H. L. (2022). A Synthesis of the Dibble et al.

343   Controlled Experiments into the Mechanics of Lithic Production. *Journal of Archaeological*

344   *Method and Theory*. https://doi.org/10.1007/s10816-022-09586-2

345   Lin, S. C., Rezek, Z., & Dibble, H. L. (2018). Experimental Design and Experimental Inference

346   in Stone Artifact Archaeology. *Journal of Archaeological Method and Theory*, *25*(3), 663–

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

348   Liu, C., Khreisheh, N., Stout, D., & Pargeter, J. (2023). Differential effects of knapping skill

349   acquisition on the cultural reproduction of Late Acheulean handaxe morphology:

350   Archaeological and experimental insights. *Journal of Archaeological Science: Reports*, *49*,

351   103974. https://doi.org/10.1016/j.jasrep.2023.103974

352   Liu, C., & Stout, D. (2023). Inferring cultural reproduction from lithic data: A critical review.

353   *Evolutionary   Anthropology: Issues,   News,   and   Reviews*,   *32*(2),   83–99.

354   https://doi.org/10.1002/evan.21964

355   Lombao, D., Guardiola, M., & Mosquera, M. (2017). Teaching to make stone tools: new

356   experimental evidence supporting a technological hypothesis for the origins of language.

357   *Scientific Reports*, *7*(1), 14394. https://doi.org/10.1038/s41598-017-14322-y

358   Mahaney, R. A. (2014). Exploring the complexity and structure of acheulean stoneknapping

359   in   relation   to   natural   language.   *PaleoAnthropology*,   *2014*,   586606.

360   https://doi.org/10.4207/PA.2014.ART90

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

362   Press.

363   Maloney, T. R., & Street, M. (2020). Hot debate: Identifying heat treatment in Australian

364   archaeology using science and modern indigenous knowledge. *Quaternary Science Reviews*,

365   *241*, 106431. https://doi.org/10.1016/j.quascirev.2020.106431

366   Martellotta, E. F., Perston, Y. L., Craft, P., Wilkins, J., & Langley, M. C. (2022). Beyond the

367   main function: An experimental study of the use of hardwood boomerangs in retouching

368   activities. *PLOS ONE*, *17*(8), e0273118. https://doi.org/10.1371/journal.pone.0273118

369   Mindermann, S., & Armstrong, S. (2018). Occam's razor is insufficient to infer the

370   preferences of irrational agents. *Proceedings of the 32nd International Conference on Neural*

371   *Information Processing Systems*, 5603–5614.

372   Mobbs, D., Wise, T., Suthana, N., Guzmán, N., Kriegeskorte, N., & Leibo, J. Z. (2021). Promises

373   and   challenges   of   human   computational   ethology.   *Neuron*,   *109*(14),   2224–2238.

374   https://doi.org/10.1016/j.neuron.2021.05.021

375   Montgomery, L. M., & Fryer, T. C. (2023). The future of archaeology is (still) community

376   collaboration. *Antiquity*, *97*(394), 795–809. https://doi.org/10.15184/aqy.2023.98

377   Moore, M. W. (2020). Hominin Stone Flaking and the Emergence of 'Top-down' Design in

378   Human   Evolution.   *Cambridge   Archaeological   Journal*,   *30*(4),   647–664.

379   https://doi.org/10.1017/S0959774320000190

380 Muller, A., Shipton, C., & Clarkson, C. (2023). The Proceduralization of Hominin Knapping

381 Skill: Memorizing Different Lithic Technologies. *Cambridge Archaeological Journal*, 1–18.

382 https://doi.org/10.1017/S0959774323000070

383 Nami, H., G. (2010). Theoretical Reflections on Experimental Archaeology and Lithic

384 Technology: Issues on Actualistic Stone Tools Analysis and Interpretation. In H. Nami G.

385 (Ed.), *Experiments and Interpretation of Traditional Technologies: Essays in Honor of Errett*

386 *Callahan* (pp. 91–168). Ediciones de Arqueologìa Contempornea.

387 Nastase, S. A., Goldstein, A., & Hasson, U. (2020). Keep it real: rethinking the primacy of

388 experimental control in cognitive neuroscience. *NeuroImage*, *222*, 117254.

389 https://doi.org/10.1016/j.neuroimage.2020.117254

390 Outram, A. K. (2008). Introduction to experimental archaeology. *World Archaeology*, *40*(1),

391 1–6. https://www.jstor.org/stable/40025310

392 Pargeter, J., Liu, C., Kilgore, M. B., Majoe, A., & Stout, D. (2023). Testing the Effect of Learning

393 Conditions and Individual Motor/Cognitive Differences on Knapping Skill Acquisition.

394 *Journal of Archaeological Method and Theory*, *30*(1), 127–171.

395 https://doi.org/10.1007/s10816-022-09592-4

396 Pfleging, J., Iovita, R., & Buchli, J. (2019). Influence of force and duration on stone tool wear:

397 results from experiments with a force-controlled robot. *Archaeological and Anthropological*

398 *Sciences*, *11*(11), 5921–5935. https://doi.org/10.1007/s12520-018-0729-0

399 Premo, L. S. (2010). Equifinality and explanation: Thoughts on the role of agent-based

400 modeling in postpositivist archaeology. In A. Costopoulos & M. W. Lake (Eds.), *Simulating*

401 *Change: Archaeology Into the Twenty-first Century* (pp. 28–37). University of Utah Press.

402 Proffitt, T., Bargalló, A., & Torre, I. de la. (2022). The Effect of Raw Material on the

403 Identification of Knapping Skill: a Case Study from Olduvai Gorge, Tanzania. *Journal of*

404 *Archaeological Method and Theory*, *29*(1), 50–82. https://doi.org/10.1007/s10816-021-

405 09511-z

406   Ranhorn, K. L., Pargeter, J., & Premo, L. S. (2020). Investigating the evolution of human

407   social learning through collaborative experimental archaeology. *Evolutionary Anthropology:*

408   *Issues, News, and Reviews*, *29*(2), 53–55. https://doi.org/10.1002/evan.21823

409   Reeves, D., Bury, R., & Robinson, D. W. (2009). Invoking occam's razor: Experimental

410   pigment processing and an hypothesis concerning emigdiano chumash rock art. *Journal of*

411   *California   and   Great   Basin   Anthropology*,   *29*(1),   59–67.

412   https://www.jstor.org/stable/27825902

413   Reeves Flores, J. (2012). *Experimental archaeology: an ethnography of its perceived value*

414   *and   impact   in   archaeological   research*   [PhD   thesis].

415   https://ore.exeter.ac.uk/repository/handle/10871/9041

416   Reynolds, P. J. (1999). The nature of experiment in archaeology. In A. Harding (Ed.),

417   *Experiment and design: Archaeological studies in honour of John Coles* (pp. 156–162). Oxbow

418   Books.

419   Roe, B. E., & Just, D. R. (2009). Internal and external validity in economics research:

420   Tradeoffs between experiments, field experiments, natural experiments, and field data.

421   *American   Journal   of   Agricultural   Economics*,   *91*(5),   1266–1271.

422   https://www.jstor.org/stable/20616293

423   Roux, V., & David, É. (2005). *Planning abilities as a dynamic perceptual-motor skill: an*

424   *actualist study of different levels of expertise involved in stone knapping* (V. Roux & B. Bril,

425   Eds.;   pp.   91–108).   McDonald   Institute   for   Archaeological   Research.

426   https://shs.hal.science/halshs-00120262

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

428   Schillinger, K., Mesoudi, A., & Lycett, S. J. (2016). Copying error, evolution, and phylogenetic

429   signal in artifactual traditions: An experimental approach using "model artifacts". *Journal of*

430   *Archaeological Science*, *70*, 23–34. https://doi.org/10.1016/j.jas.2016.04.013

431   Schmidt, S. C., & Marwick, B. (2020). Tool-Driven Revolutions in Archaeological Science.

432   *Journal of Computer Applications in Archaeology*, *3*(1), 1832.

433   https://doi.org/10.5334/jcaa.29

434   Shamay-Tsoory, S. G., & Mendelsohn, A. (2019). Real-Life Neuroscience: An Ecological

435   Approach to Brain and Behavior Research. *Perspectives on Psychological Science*, *14*(5),

436   841–859. https://doi.org/10.1177/1745691619856350

437   Simon, D. A., Gordon, A. S., Steiger, L., & Gilmore, R. O. (2015). *Databrary: Enabling sharing*

438   *and reuse of research video.* 279280. https://doi.org/10.1145/2756406.2756951

439   Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic Stimuli in Neuroscience:

440   Critically Acclaimed. *Trends in Cognitive Sciences*, *23*(8), 699–714.

441   https://doi.org/10.1016/j.tics.2019.05.004

442   Stout, D. (2002). Skill and cognition in stone tool production: An ethnographic case study

443   from irian jaya. *Current Anthropology*, *43*(5), 693–722. https://doi.org/10.1086/342638

444   Stout, D. (2021). The cognitive science of technology. *Trends in Cognitive Sciences*, *25*(11),

445   964–977. https://doi.org/10.1016/j.tics.2021.07.005

446   Stout, D., Chaminade, T., Apel, J., Shafti, A., & Faisal, A. A. (2021). The measurement,

447   evolution, and neural representation of action grammars of human behavior. *Scientific*

448   *Reports*, *11*(1). https://doi.org/10.1038/s41598-021-92992-5

449   Syed, M., & McLean, K. C. (2022). Disentangling paradigm and method can help bring

450   qualitative research to post-positivist psychology and address the generalizability crisis.

451   *Behavioral and Brain Sciences*, *45*, e32. https://doi.org/10.1017/S0140525X21000431

452   Timbrell, L. (2023). A Collaborative Model for Lithic Shape Digitization in Museum Settings.

453   *Lithic Technology*, *48*(1), 31–42. https://doi.org/10.1080/01977261.2022.2092299

454   Varela, F. J., Thompson, E., & Rosch, E. (2017). *The Embodied Mind: Cognitive Science and*

455   *Human Experience* (revised edition). The MIT Press.

456   Weger, U. W., Wagemann, J., & Tewes, C. (2019). Editorial: The challenges and opportunities
457   of introspection in psychology: Theory and method. *Frontiers in Psychology*, *10*.
458   https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02196

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

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

463   Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A.,
464   Blomberg, N., Boiten, J.-W., Silva Santos, L. B. da, Bourne, P. E., Bouwman, J., Brookes, A. J.,
465   Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B.
466   (2016). The FAIR Guiding Principles for scientific data management and stewardship.
467   *Scientific Data*, *3*(1), 160018. https://doi.org/10.1038/sdata.2016.18

468   Yarkoni, T. (2022). The generalizability crisis. *Behavioral and Brain Sciences*, *45*, e1.
469   https://doi.org/10.1017/S0140525X20001685

1
2
3
4
5
6
7
8
9
10
11
12
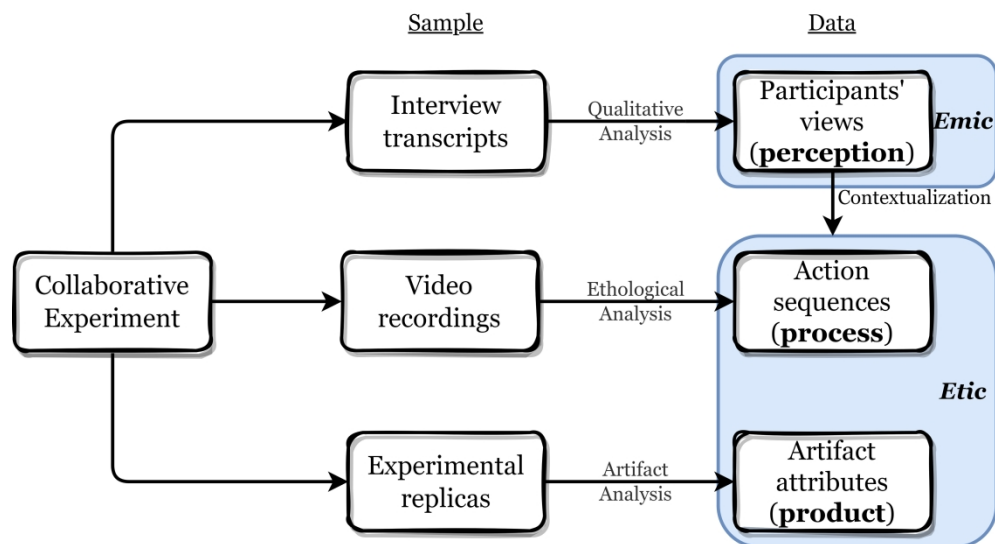13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
...

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

395x227mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
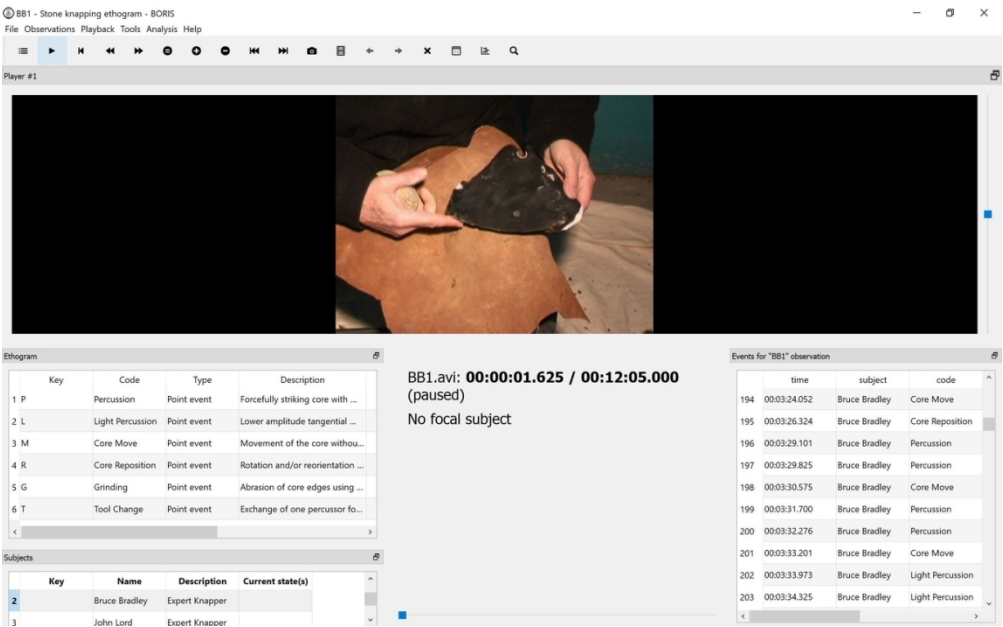51
52
53
54
55
56
57
58
59
60

Fig.2. An example of coding a handaxe knapping session using the BORIS software.

115x71mm (300 x 300 DPI)