






## Three levels of reproducible workflow remove barriers for archaeologists and increase accessibility

**Ben Marwick**  based on peer reviews by **Cyler Conrad**, **Cheng Liu** , **Sam Leggett**  and **Lisa Lodwick** ?

Emma Karoune and Esther Plomp (2022) Removing Barriers to Reproducible Research in Archaeology. Missing preprint\_server, ver. 5, peer-reviewed and recommended by Peer Community in Archaeology. <https://doi.org/10.5281/zenodo.7320029>

Submitted: 07 June 2022, Recommended: 21 November 2022

### Cite this recommendation as:

Marwick, B. (2022) Three levels of reproducible workflow remove barriers for archaeologists and increase accessibility. *Peer Community in Archaeology*, 100022. <https://doi.org/10.24072/pci.archaeo.100022>

Published: 21 November 2022

Copyright: This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

---

Over the last decade, a small but growing community of archaeologists, from a diversity of intellectual and demographic backgrounds, have been striving for computational reproducibility in their published research. In their survey of the accomplishments of this thriving community, Emma Karoune and Esther Plomp (2022) analyzed the wide variety of approaches researchers have taken to enhance the reproducibility of their research. A key contribution of this paper is their excellent synthesis of diverse approaches into three levels of increasing complexity. This is helpful because it provides multiple entry points for researchers new to the challenge of fortifying their research. Many researchers assume that computational reproducibility is only achievable if they have a high degree of technical skill with computers, or is only necessary if their work is very computationally intensive. Karoune and Plomp give three compelling reasons why reproducibility is important for all archaeological research, and through their three levels they demonstrate that how these levels can be accomplished with basic, non-specialized computer skills and widely used free software. They showcase exemplary work from a variety of archaeologists to show how practical and achievable reproducible research is for all archaeologists. They advocate for archaeologists to use the most widely used and supported tools and services to support their reproducible research, such as the R and Python programming languages for data analysis, and Git and GitHub for collaboration.

This paper, with its extensive appendix including thoughtful responses to frequently asked questions about reproducible research in archaeology, is likely to have a wide reach and influence, beyond previous works on this topic that have largely focused on technical details. Karoune and Plomp have provided the on-ramp for a generation of archaeologists who will find their questions about reproducible research answered here. They

will also find an agreeable entry point to reproducible research in one of the three levels described by the authors. Will every archaeologist embrace this way of working? Should they? The work of Leonelli (2018) can help us anticipate the answers to these questions. Leonelli asks where are the limits to reproducibility, and how do the characteristics of different ways of knowing affect the desirability of reproducibility? Leonelli's work invites us to consider that there will be archaeologists coming from different epistemic cultures for whom the motivations presented by Karoune and Plomp will not resonate. For example, archaeologists engaged in mostly hermeneutical social science and humanities research, who do little or no quantitative analysis and statistics, are unlikely to see reproducibility as meaningful or desirable for their work. We can describe these researchers as working in interpretative or constructivist epistemic cultures. In these cultures, the particulars of how an individual researcher engages with their subject are exclusive and unique, and they would argue it cannot be fully captured or shared in a meaningful way (Elman and Kapiszewski 2017). Here, knowledge is situational, emerging from a specific, once-off combination of people and circumstances. One example in archaeology is the *chaîne opératoire* approach of stone artefact analysis, which Monnier and Missal (2014:61) describe as "based upon the analyst's experience and intuition, and it is not replicable, nor quantifiable". To make sense of this example we can draw on Galison's (1997) concept of 'image traditions' and 'logic traditions'. An image tradition is a way of knowing that is qualitative, based on composing narratives from drawings and photographs. A logic tradition is based on the use of instruments and statistical methods to collect standardised quantitative data. *Chaîne opératoire* approaches fall into the image tradition, along with many other ways of working in archaeology that do not generate numbers or use them to support claims about the past. Archaeologists working in a logic tradition will find reproducible research to be more meaningful than those working in an image tradition.

We should be mindful not to claim that one epistemic culture is superior to another because reproducibility is not meaningful or attainable for researchers in one culture. Such a claim would threaten the plurality that is essential for the reliability of scientific knowledge (Massimi 2022). Instead we should identify those communities in archaeology where reproducible research is both meaningful and attainable, but has not yet been widely embraced. That is the where the most beneficial effects can be expected. According to Leonelli's (2018) framework, we can recognise these communities by a few basic characteristics. For example: they are doing computationally intensive archaeology, such as using or writing software to collect, simulate, analyse or visualise data; they are doing experimental archaeology; or they are making knowledge claims that are supported by tables of numeric data and data visualisations. Archaeologists whose work shares one or more of these characteristics will find the guidance provided here by Karoune and Plomp to be highly instructive and relevant, and stand the most to benefit from it.

But it is not only individual archaeological scientists that have potential to benefit from how Karoune and Plomp have lowered the barriers to reproducible research. An especially important implication of this paper is that by lowering the barriers to reproducible research, Karoune and Plomp help us all to lower barriers to participation in archaeology in general. Documenting our research transparently, and sharing our materials (such as data and code and so on) openly, can profoundly change how others can participate in archaeology. By doing this, we are enabling students and researchers elsewhere, for example in low and middle income locations, to use our materials in their teaching and learning. Other researchers and students can apply our methods to their data, and combine their data with ours to achieve syntheses beyond what a single project can do. Similarly, for archaeologists working with local, descendant or marginalized communities, the tools of reproducible research are vital for enabling community members to have full access to the archaeological process, and thus reproducibility may be considered a necessity for decolonising the discipline. Karoune and Plomp present the CARE principles (Carroll et al. 2020) to guide archaeologists in ensuring community control of data so that reproducibility can be ethically accomplished with community safety and well-being as a priority. This may have a profoundly positive impact on the demographics of archaeology, as it lowers the barriers of meaningful participation by people far beyond our immediate groups of collaborators.

Making archaeology more accessible is of critical importance in stemming the negative social impacts of

pseudoarchaeologists, who often claim that archaeologists actively suppress the truth of the archaeological record through secrecy, elitism, and exclusiveness. The harm in this is twofold. First, that pseudoarchaeology typically erases Indigenous heritage by claiming that their past achievements were due to an ancient, extinct advanced civilization, not Indigenous people. These claims are often adopted by white supremacists to support racist and antisemitic conspiracy theories (Turner and Turner 2021), which sometimes leads to prejudice, physical violence, radicalization and extremism. A second type of harm that can come from claims of secrecy and elitism is it drains public trust in experts, leading to science denial. Not only trust in archaeologists, but trust in many kinds of experts, including those working on urgent contemporary issues such as public health and climate change. Karoune and Plomp's work is important here because it provides a practical and affordable pathway for archaeologists to fight claims of secrecy and elitism by sharing their work in ways that make it possible for non-academics to inspect the analyses and logic in detail. Claims of secrecy and elitism can be easily countered by openness, transparency and reproducibility by archaeologists. This is not only useful for tackling pseudoarchaeologists, but also in enacting an ethic of care, framing members of the public as people that not only *care about* archaeology as part of humanity's shared heritage, but also *care for* the construction of reliable interpretations of the archaeological record to provide secure and authentic foundations for their social identities and relationships (Wylie et al 2018; de la Bellacasa 2011). By striving for reproducible research in the way described by Karoune and Plomp, we are practicing a kind of reciprocal care among ourselves as archaeologists, and between archaeologists and members of the public as two communities who care about the human past.

### **References:**

- Karoune, E., and Plomp, E. (2022). Removing Barriers to Reproducible Research in Archaeology. Zenodo, 7320029, ver. 5 peer-reviewed and recommended by Peer Community in Archaeology. <https://doi.org/10.5281/zenodo.7320029>
- de la Bellacasa, M. P. (2011). Matters of care in technoscience: Assembling neglected things. *Social Studies of Science*, 41(1), 85–106. <https://doi.org/10.1177/0306312710380301>
- 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., and Hudson, M. (2020). The CARE Principles for Indigenous Data Governance. *Data Science Journal*, 19(1), Article 1. <https://doi.org/10.5334/dsj-2020-043>
- Elman, C., and Kapiszewski, D. (2017). Benefits and Challenges of Making Qualitative Research More Transparent. *Inside Higher Ed* 2017, <https://www.insidehighered.com/blogs/rethinking-research/benefits-and-challenges-making-qualitative-research-more-transparent> (accessed 21 Oct, 2022).
- Galison, P. (1997). *Image and logic: a material culture of microphysics*. Chicago (IL): University of Chicago Press.
- Leonelli, S. (2018). Re-Thinking Reproducibility as a Criterion for Research Quality [preprint]. Available online: <http://philsci-archive.pitt.edu/id/eprint/14352> (Accessed 21 Oct 2022).
- Massimi, M. (2022). *Perspectival realism*. Oxford University Press.
- Monnier, G. F., and Kele M.. "Another Mousterian debate? Bordian facies, chaîne opératoire technocomplexes, and patterns of lithic variability in the western European Middle and Upper Pleistocene." *Quaternary International* 350 (2014): 59-83. <https://doi.org/10.1016/j.quaint.2014.06.053>

Turner, D. D., and Turner, M. I. (2021). "I'm Not Saying It Was Aliens": An Archaeological and Philosophical Analysis of a Conspiracy Theory. In A. Killin and S. Allen-Hermanson (Eds.), *Explorations in Archaeology and Philosophy* (pp. 7–24). Springer International Publishing.

[https://doi.org/10.1007/978-3-030-61052-4\\_2](https://doi.org/10.1007/978-3-030-61052-4_2)

Wylie, C., Neeley, K., and Ferguson, S. (2018). Beyond Technological Literacy: Open Data as Active Democratic Engagement? *Digital Culture & Society*, 4(2), 157–182.

<https://doi.org/10.14361/dcs-2018-0209>

## Reviews

### Evaluation round #2

DOI or URL of the preprint: <https://doi.org/10.5281/zenodo.7320029>

Version of the preprint: v2

### Authors' reply, 27 October 2022

Hi Ben,

Thank you for taking the time to edit our manuscript. We have accepted all your comments apart from a few - mostly the we want to keep figure 1. We have also address the language by making the main article passive voice and then the appendix more of a active voice by taking yor advice on changing 'I' to 'you'. We have also added in references to stregthen the argument in the intro section and also the part about stand alone articles not being reproducible.

We have uploaded a marked up version here so you can see the changes and our replies to your comments.

The clean version 4 can be found here: <https://doi.org/10.5281/zenodo.7256954>

Thanks for your help,

Emma Karoune and Esther Plomp.

[Download tracked changes file](#)

### Decision by **Ben Marwick** , posted 24 October 2022

#### **Three levels of reproducible workflow remove barriers for all archaeologists**

Thank you Emma and Esther for your thoughtful responses and diligent revisions. I have made editorial changes to the text, starting with that MS Word document, and using track changes. In brief, I have made the main focus of the paper the three levels, and moved all the Q&A to the appendix. I think this gives the paper a more coherent and compelling logical structure, one that is consistent with what most readers will be expecting. Most readers are expecting a sustained, contextualised argument in a journal article, and I believe they come to PCI expecting to find pieces that closely resemble journal articles. In this case the main argument of this paper is that approaches to reproducible research can be organised into three levels. I think this is the most original and creative contribution in this paper, and deserved a stronger focus.

My perspective is that the Q&A content, as it is currently written, detracts from the main argument and was not effective at contextualising the main argument. Some of the Q&A is just a list of links, which is fine for a workshop handout, but I believe inconsistent with most people's expectations of a journal article. I agree that

the Q&A text is relevant to the paper, so I moved it all into the appendix. I've also divided the appendix into two appendices: the Q&A and the glossary.

I've edited the main text of the paper to make it consistent with my perspective on 'high scientific quality' because 'PCI Archaeology recommends only preprints of high scientific quality that are methodologically and ethically sound.' ([https://archaeo.peercommunityin.org/help/guide\\_for\\_authors](https://archaeo.peercommunityin.org/help/guide_for_authors)). Editing this paper is challenging because the writing frequently switches between passive third person to active second person. The active second person voice is very rare in journal articles, and I'm concerned that many readers will conflate its presence here with low quality scholarship. This is because readers are not accustomed to being so directly addressed in journal articles, and I believe some will find it a bit off-putting with the gap between the author and reader so small. To be clear, I think this paper contains high quality scholarship, and I want it to have an extensive readership and impact. I think one way we can support that is to satisfy readers' basic expectations. I believe they come to PCI Archaeology expecting to read journal articles, so we should tailor our writing to meet those expectations and follow some of the conventions of journal article writing. Otherwise the reader will question the credibility and reliability of what they are reading. So that's my main motivation for editing the main text.

I've only very lightly edited the appendixes, since I think readers have different expectations of those. I found the text formatting, e.g. size, bold and italics, inconsistently applied throughout, which gives the reading a feeling of disorder. I encourage you to take a very systemic approach to using those text decorations. The shifting uses of "I", "you", and "we" is also jarring throughout the appendix. I think this could be easily fixed by replacing "I" with "you" throughout.

If you are ok with my edits and can submit a version with all the tracked changes accepted and questions responded to, I'll mark it as 'recommended' and prepare a note to appear on PCI.

Thank you,

Ben [Download recommender's annotations](#)

## Evaluation round #1

DOI or URL of the preprint: <https://doi.org/10.5281/zenodo.7256954>

### Authors' reply, 22 October 2022

Dear Reviewers,

Thank you so much for the positive responses to our article. We were delighted that it was received so well and the review comments were extremely helpful in improving the article.

We have addressed most of the comments by making additions to the paper. However, some of the comments we felt would be large additions and we therefore decided not to incorporate them. This was really due to the fact that the article is already very long.

Please find the new version (v3) of the paper at this doi: <https://doi.org/10.5281/zenodo.7239193>

Best wishes,

Emma Karoune and Esther Plomp.

[Download author's reply](#)

[Download tracked changes file](#)

Decision by [Ben Marwick](#) , posted 19 August 2022, validated 25 October 2022

Dear Dr Karoune and Dr Plomp,

Thank you for submitting your pre-print for review, and for providing an opportunity for a robust and stimulating discussion about reproducibility in archaeology. I have been so inspired by similar discussions in other disciplines (e.g. handy guides for beginners such as Alston and Rick 2020 in ecology and revealing surveys of barriers such as Stodden 2010 in computer science), and I believe that essays such as the one you have written will similarly inspire and guide many archaeologists to improve the reproducibility of their work. An especially motivating detail that you mention is the importance of reproducible research for supporting sustainability, inclusiveness, and equitable access to participating in archaeological research.

Thanks also to our four reviewers, who are some of the most skilled and experienced scholars on this topic. It's an honour to have input from these researchers who have pioneered reproducibility in many areas of archaeology, and whose own compendia of code and data should be the among the first things junior scholars seek out as excellent examples of how to do this (e.g. Conrad et al 2016; 2021; Leggett 2021; 2022; Lodwick 2019).

Dr Karoune and Dr Plomp, please do carefully study the thoughtful reviews and consider editing and expanding your paper as they recommend. There are many excellent suggestions that will greatly help in upgrading your pre-print from something of a workshop handout, as it is currently, to a substantial manuscript with broad relevance to archaeologists around the world that helps to advance reproducibility in archaeology. As you do your revisions, I hope you might be able to draw relevant and diverse examples of reproducible research in this list of 250+ archaeology articles spanning 10 years that include R code and data.

A technical note: many of the resources you cite are websites without persistent identifiers, and so there is a danger of link-rot in your paper that will be frustrating for future readers. For an ephemeral workshop handout, this is expected, but for a scholarly publication I think we should invest some effort into insuring against the risk of link-rot to make the paper useful to readers long into the future. I recommend including only the most relevant and stable links in your paper, and removing those that are already out of date (I found a few that reference outdated content) or less relevant to your central claims. Then I recommend, as much as it practical, replacing links in your text with traditional in-text citations, following a widely used style such as APA. This will help readers to find the websites if there are minor changes to the URLs, which happen often. Additionally I recommend including in the reference to each website an archive URL from a service such as <https://perma.cc/> or <https://web.archive.org/> Then readers can still access the content even after the original website has gone.

#### References cited

Alston, J. M., and Rick, J. A.. 2020. A Beginner's Guide to Conducting Reproducible Research. *Bull Ecol Soc Am* 102( 2):e01801. <https://doi.org/10.1002/bes2.1801>

Conrad, C., et al. (2021). Re-Evaluating Pleistocene–Holocene Occupation of Cave Sites in North-West Thailand: New Radiocarbon and Luminescence Dating. *Antiquity* <https://doi.org/10.15184/aqy.2021.44>

Conrad, C., et al. (2016). Paleoeology and Forager Subsistence Strategies During the Pleistocene-Holocene Transition: A Reinvestigation of the Zooarchaeological Assemblage from Spirit Cave, Mae Hong Son Province, Thailand. *Asian Perspectives* 55(1). <https://www.jstor.org/stable/26357698>

Leggett, S. (2022). A Hierarchical Meta-Analytical Approach to Western European Dietary Transitions in the First Millennium AD. *European Journal of Archaeology*, 1-21. <https://doi.org/10.1017/ea.2022.23>

Leggett, S. (2021). Migration and cultural integration in the early medieval cemetery of Finglesham, Kent, through stable isotopes. *Archaeol Anthropol Sci* 13, 1. <https://doi.org/10.1007/s12520-021-01429-7>

Lodwick, L., 2019. Sowing the Seeds of Future Research: Data Sharing, Citation and Reuse in Archaeobotany. *Open Quaternary*, 5(1), p.7. <http://doi.org/10.5334/oq.62>

Stodden, Victoria, *The Scientific Method in Practice: Reproducibility in the Computational Sciences* (February 9, 2010). MIT Sloan Research Paper No. 4773-10, <http://doi.org/10.2139/ssrn.1550193>

## Reviewed by Lisa Lodwick, 06 July 2022

This is a very useful article setting out accessible routes to reproducible research in archaeology. The key merits of this article are the setting out of definitions of, barriers to, and solutions for reproducible research in archaeology in an accessible, richly resourced and easily understandable format. The Appendix and FAQs are a rich resource for those looking to adopt open science practices in their research, and the profusion of hyperlinks are welcome. The article text itself is at times a touch brief in alluding to much bigger debates. The high-level review style aids accessibility, but there is perhaps a lack of engagement in longer term debates about the structure and philosophy of archaeological research, and initiatives around databases, data-sharing and LOD. I have suggested some small additions of archaeological examples to keep the archaeological reader engaged, and I consider that the richness of information and clarity of the article make it a beneficial and useful article for many archaeologists. The article is well cited, and links to a range of research products. The figures and tables, language and structure are all clear. My comments below relate to minor rewording to ensure clarity, and the addition of more archaeological examples to aid the goal of the article.

### Abstract

- Minor language point in sentence 3 – are these barriers or requirements? If using barriers, as in the title, then I would suggest rewording this list as the barrier ‘skill level of researchers’ or ‘software and infrastructure availability’.

### Introduction

- This section clearly outlines the timeliness of the article.
- Some more content on the motivation for this paper would be useful.
- I would suggest adding a ‘hook’ on why reproducibility for archaeology matters – why should an archaeologist invest the time to read this article – what will making their research more reproducible do for them? Some of these aspects are raised later in the article E.g. archaeology as unique observations, but I would suggest including a few here to keep the readers interest. I would suggest considering whether the profusion of micro-specialisms with distinct methodologies and data standards, plus the blurred disciplinary position may be other reasons why reproducible practices would be of benefit to archaeologists.

- I am supportive of this small step approach outlined in this article – could the authors give an example of another discipline or field where such small step individual learning approaches have worked? It would be useful to briefly mention other pathways towards reproducibility, such as stringent peer reviewing of study reproducibility by journals, or inclusion of reproducibility within under- and post-graduate training.

### Why is reproducible archaeology important?

- Paragraph 4 – I would question why rescue or commercial work alone are highlighted for their time limited nature – Research excavations are also time-limited and destructive.

- Paragraph 5 – The authors could also consider money limitations for scientific analysis such as machine time, consumables, technician time etc.

- Line 177 sentence beginning “Consequently...” – I suggest a clarification of this sentence– make clear that you are calling for reproducible research to enable reassessment from the point of the dataset, rather than the point of the material assemblage.

- Para beginning line 180 – I’m unclear what Transparent recording is referring to – is this the recording of the original data, or of the analysis undertaken? An example here would be useful.

- Spelling out the CARE principles would be useful here – FAIR principles now appear fairly widely known and used in archaeology, less so with the CARE principles. I see you do this in more detail below, I would briefly allude to them in this section.

### What does reproducible research look like?

- Line 216 , the sentence beginning “Large meta-analysis” – I suggest rewording to improve clarity. Is it that both the large meta-analysis studies and studies that want to reuse the same methods need computational reproducibility?

- Para beginning Line 224- perhaps introduce the concept that the data code and methods are files in the previous paragraph.
- Line 258 – can you give examples of proprietary software – SPSS, Excel etc.
- Table 1 is useful – could you add some more examples dealing with different forms of archaeological data? I think they would be very useful to readers.
- Line 296 – perhaps highlight that the Analysis output file is different to a graphical depiction of these analysis results, which is the much more ubiquitous feature of archaeological publications.
- Line 329 – Again, some archaeological examples of research compendium would be really useful here.
- Line 447 – with regards to sensitive location data, the PAS in the UK provides a good example of dealing with sensitive find spot location (<https://finds.org.uk/help/database/topic/id/10>)
- Line 462 – Another scenario which you might want to consider is when researchers are using analysis code shared within a lab group – it may be written in R, but developed by someone else, and passed on between researchers, so not the researchers to share.
- Line 472 – I would suggest mentioning citation of data sets as an encouragement to data sharing.

#### 5. Join a community or association

- I would suggest the addition of the UK Reproducibility Network, which has local nodes across the UK <https://www.ukrn.org/>

#### Conclusion

Currently, this reads like the end of a help document – I would suggest reiterating points from the first half of the article on the importance of, and barriers to, the current adoption of reproducible research in archaeology.

The FAQ is very useful, especially the section on preregistration.

### Reviewed by **Sam Leggett** , 06 July 2022

This article is a fantastic piece which explains how and why reproducible and open research in archaeology is attainable and important for moving the discipline forward. I hope there are other follow-up pieces in the works to provide more detailed information for researchers on specific aspects of mentioned within. I really enjoyed reading this piece and its detailed appendices, it is a long overdue article for our field, and will have a lasting resonance across all the archaeological sub-disciplines. I will certainly be adding it to my student resource lists and taking on board many of the suggestions and resources suggested in here for my own research.

Overall, I thoroughly recommend this paper for publication, it's extremely well written, timely, covers important ground about the future of open workflows in our discipline and will serve as a foundational guide moving forward. It has lots of great resources and examples which are extremely helpful to researchers at various stages of their reproducibility and open workflow journeys, and answers many frequently asked questions from colleagues I've seen time and again. Below are my more detailed comments and suggestions, all of which are minor.

The extensive appendices are to be particularly commended as they detail commonly used terms in Open Science and reproducibility literature which are not accessible to the uninitiated. The glossary and Q&A sections are especially helpful, however as someone who is familiar with many of these terms I'm not sure I am best placed to suggest other areas which other colleagues might want answered or definitions of, so I hope people do interact with this Open format and leave queries for the authors to keep the conversation going.



There are a few suggestions I have for additions to the glossary – binders and containers I think could use a little further explanation for those who aren't familiar with coding and that side of software. Along with R and Python, Git/Github should also warrant an entry as they're mentioned in text for version control. 'Protocol' as a term should also be added, again I think those in more wet lab based archaeological science will be familiar, but it is a term less familiar to others who likely use their own protocols but might not label them as such. I would also recommend that the glossary be more heavily signposted from the start of the article as it is a great key resource you've created and needs to be flagged early on for reference throughout.

Figure 1 whilst nice I think isn't strictly necessary and doesn't add substantially to the text so could be left out.

The differences between replication, reproducibility, robustness, and generalisability are explained well in text but could do with further clarity in Figure 2, especially generalisable as in reality if you got completely different results with different data and different analysis (as shown in Figure 2), it is unlikely you would come up with the same generalizations about a phenomenon. Similarly robust could do with clarification in Fig 2 for the same reasons. But again I stress this is made clear in text, but to limit misinterpretation on reuse of Figure 2 I'd edit it slightly.

Figure 3 needs an explicit citation or source link.

I suggest adding a little more detail around lines 131-133 about the differences described, and the Marwick 2020b reference does not align with how that paper is listed in the bibliography, please cross-check all references before publication as I think a few others with "a" and "b" in text have also slipped through.

The definition of the discipline on line 136, particularly as a scientific study is hotly debated, and I worry that defining it as only that may not remove barriers to all archaeologists working more openly but could put some back up, especially for those who see working reproducibly and openly as something for those on the more laboratory-based side of things. Problematising this more or offering up a broader definition of the field is advised.

The point about validation in line 199-200 is such an important and poignant one, that I feel warrants further explanation or signposting to a future article, you left me wanting more detail here, and I definitely agree with you that it extremely important and underrated.

Where you discuss large meta-analyses in lines c. 215-220 it would be great to have examples of these kinds of studies like you do for other parts of reproducible research later in the article to showcase how data can be collated and re-used.

Table 1 and the steps from line 280 onwards with the examples are fantastic! I particularly like the inclusion of various freely available and easy to use websites and software such as Google Docs alongside the more advanced tools. With the raw data file formats, csv is the only option mentioned in method 1 – perhaps include recommendations for non-tabulated data such as image files, and best practice/formats for other data types. Another suggestion to perhaps improve the uptake of your steps for reproducibility would be to expand Table 1

and parts of the text into a flowchart, checklist, or template(s) of steps for researchers to follow as a workflow as an accessible route for helping to set up either a project from the beginning or ensure your work meets as many steps as possible while preparing for publication.

The “confronting your barriers” section is a great idea and I really commend the authors for this section, especially the inclusion and consideration of CARE principles alongside the more widely touted FAIR principles. Whilst I love the idea of using synthetic data suggested in lines 434-441 I think this is a huge conceptual and training hurdle for many, and so might not be very accessible or executable in a lot of cases. Therefore, any additional resources the authors can point people to here would be a great help for how to go about creating synthetic data that meets replicability criteria.

I would also suggest adding “RLadies” to the list of communities and associations in lines 518-528.

In the appendix under the metadata standards, I would also include those widely used in the journal “Ecology” which have been widely applied in aspects of environmental archaeology, especially isotopic data collation - Michener, William K., James W. Brunt, John J. Helly, Thomas B. Kirchner, and Susan G. Stafford. ‘Nongeospatial Metadata for the Ecological Sciences’. *Ecological Applications* 7, no. 1 (1997): 330–42. [https://doi.org/10.1890/1051-0761\(1997\)007\[0330:NMFTES\]2.0.CO;2](https://doi.org/10.1890/1051-0761(1997)007[0330:NMFTES]2.0.CO;2). Also here: <https://www.esa.org/wpcontent/uploads/2022/05/ESA-Data-Paper-Guidelines.pdf>

The pre-registration section is such a welcome addition – thank you! I think this is something we really do not take advantage of enough in archaeology.

In the section “my supervisor won’t let me work reproducibly...” I’d also add in links to the Wellcome Trust funding guidelines on working openly here as they are a fantastic resource with lots of signposting to further resources. To make this more globally inclusive perhaps also link to other research councils like those in Australia (<https://www.arc.gov.au/aboutarc/program-policies/open-access-policy>), the NSF in the United States (<https://www.nsf.gov/pubs/2016/nsf16009/nsf16009.jsp#q1><https://www.nsf.gov/pubs/2016/nsf16009/nsf16009.jsp#q1>) and this policy draft for India (<https://openaccessindia.org/national-open-access-policy-of-india-draft-ver-3/><https://openaccessindia.org/national-open-access-policy-of-india-draft-ver-3/>).

My final suggestion would be that your last sentence referring readers to the Kansa et al. paper could be duplicated earlier in the main text and more frequently – I know you cite it regularly as it’s such a key paper but hammering it home as you do here earlier would be beneficial as some people may not look at the very end of the appendices.

Thank you for all of your hard work and dedication compiling this fantastic paper and its thorough appendices!

**[Download the review](#)**

## Reviewed by [Cyler Conrad](#), 06 July 2022

Karoune and Plomp present an insightful, functional, and significant discussion of reproducible research in archaeology within this manuscript. Their stated goal, to “introduce reproducible research in an understandable manner so that archaeologists can learn where and how to start improving the reproducibility of their research,” is certainly achieved with examples, workflows, clear definitions, and more. A caveat to my review: I do not consider myself an expert in the philosophy of reproducibility in archaeology, I see myself instead as a practitioner of my own niche understanding of how best to “do” reproducible research (e.g., <https://github.com/cylerc>). I share this because the strength of this manuscript is that archaeologists who are new, or experienced, practitioners in reproducibility will undoubtedly find helpful guidance within Karoune’s and Plomp’s manuscript – I certainly have, and I thank the authors for their exhaustive efforts in making these concepts accessible to us all within archaeology (and perhaps elsewhere!). Readers will take away an appreciation for why reproducibility matters and how to accomplish reproducible research.

The manuscript includes an introduction with key background (i.e., what is reproducibility, why is it important, etc.) and excellent figures/illustrations. In fact, the illustrations throughout this entire manuscript are impressive. The authors clearly understand the necessary intersection between text and visual aids for the greatest dissemination of concepts (reproducibility of ideas within a reproducibility manuscript). Karoune and Plomp then provide three examples of reproducible workflows, and discussions on barriers in reproducibility, training, resources, definitions, and an appendix of frequently asked questions and additional resources.

I have virtual no substantive constructive comments on the manuscript in its current form. As I previously mentioned, I gained new insights into the process (and capabilities) of reproducibility in archaeology while conducting this review. There are only a few minor thoughts that the authors may be interested in considering:

Line 180-189: I certainly agree that we need to move away from the sole ownership of research kept on our local computers that only benefit ourselves and a few other researchers. That seems to be a critical aspect of reproducibility, or really the *success* of reproducibility in archaeology. This also brought up an issue that I think about often which is the ability to share and practice reproducibility within non-academic (not a great term for this distinction) settings. I’m thinking particularly about State/Federal/Private agency and company archaeological records. For example, how do archaeologists practice reproducibility in their work when their research occurs within an agency that is consulting with Indigenous Nations, States, Federal agencies, or some combination of the above? Their research – regardless of the scale – still contributes to the archaeological body of knowledge, but there may be regulatory or proprietary reasons that data, concepts, results, even ideas, cannot be shared. There does not seem to be an easy way to currently manage this “grey” literature and reproducibility framework within archaeology except on a case-by-case basis. tDAR is a logical location where these types of data are currently curated, but I suspect a struggle is still the ability to have transparent reproducibility (e.g., tDAR might curate a record without easy access to the record itself). This is all just something to consider. It is an ongoing challenge.

Line 224-235: This is a key section, and I would recommend adding in a brief mention of the ability to use university repositories in addition to Zenodo, OSF, Figshare, and more. I recognize that in some cases university repositories have limitations, and that these limitations may also be present in non-university repositories. For example, long-term preservation of digital data, curation of servers, etc., but I suspect that in some cases archaeologists beginning their reproducibility journey might find helpful resources and support within their university system. Arizona State University has a non-exhaustive list of some university open access repositories, here: <https://libguides.asu.edu/openaccessresources/repositorieshttps://libguides.asu.edu/openaccessresources/repositories>

A final note relates to the concept of “full reproducibility” (Line ~212, Figure 4) in archaeology. Something that I have found disconcerting in our modern world of archaeology is exactly *how* research projects are created. I think this is as critical to research as how research projects occur in reproducible systems. This perhaps relates to “ideas”, “discussions”, “consultation”, “questions” under the research iceberg. I wish there were mechanisms in place to exhaustively document how research begins – or a sense of self-reporting responsibility to document how research began. Did a published paper spur a new idea? A tweet? A conference presentation? A conversation at a conference? An overheard conversation at a conference? You can see where I’m going with this...there are so many possible sources of inspiration for research in archaeology, and an equally large number of routes to accomplish that research (e.g., fully funded, and transparent research proposals vs. behind-the-scenes lab analyses through colleagues, etc.) but in very few instances are those processes made entirely clear in final products. I hope that as reproducibility in archaeology continues to evolve and take prominence that there will be a shift towards making the formation of research projects transparent and open as well. To me, this would create the ultimate cycle of reproducibility by identifying the underground roots of projects, not just the tree that grows once the roots are established (using a similar metaphor to the research iceberg). However, I’m also aware that once open, transparent, and reproducible archaeology begins to approach this aspect of the research process, there will likely be a needed evaluation of how archaeological ethics facilitate (or not) those specific research projects.

[Download the review](#)

Reviewed by [Cheng Liu](#) , 22 June 2022

This well-written manuscript provides a convincing narrative of the importance of reproducible archaeological research and a brief how-to guide for beginners in the appendix. In the meantime, it also presents a balanced and fair view of the ethical concerns behind data sharing. In short, I highly recommend this paper to be published in a peer-reviewed venue in the future and hope it can persuade more archaeologists to value and practice open and reproducible research. Below are just a few minor comments to be addressed.

**Minor Comments:**

line 400: Mukurtu (<https://mukurtu.org/about/>) is a possible data management and sharing platform when working with indigenous communities. Their code is fully available on GitHub but setting up a server might involve some cost. One example using the Murkurtu platform is the Rowasu’u project (<https://rowasu.org/about>).

line 472: Related to the scooping issue, it is very common for archaeologists to claim “data available on request” at the end of the manuscript to get away because of the mandatory data availability statement of journals. However, according to a recent study conducted by Gabelica et al. (2022), only the authors of 122 (6.8%) out of 1792 manuscripts actually responded to their request for data sharing and provided the corresponding data. I think this is a strong piece of evidence of why “data available on request” is not enough.

line 483: “more easy” should be “easier”

line 495: My personal recommendation here is a Coursera course called Reproducible Templates for Analysis and Dissemination (<https://www.coursera.org/learn/reproducible-templates-analysis>), covering the basics of Git and Rmarkdown. Everyone can enroll in the course for free, but the exercise, which is rather unnecessary, requires payment.

line 541: One more thing I would recommend here is R style guides like Google’s R Style Guide (<https://google.github.io/styleguide/Rguide.html>) and ISAAK’s R Style Guide (<https://gitlab.com/ISAAKiel/StyleGuide>). Style guide can be useful, particularly for beginners, because one psychological barrier to reproducible research is that people are worrying if their codes are too ad-hoc, messy, and inconsistent to be shared with anyone. These style guides can help beginners foster good habits of programming. Some other archaeology-related resources include general archaeological sciences using R (<https://benmarwick.github.io/How-To-Do-Archaeological-Science-Using-R/>), archaeological network analysis (<https://book.archnetworks.net/index.html>), and Marwick’s compiled

list of archaeological papers including R codes (<https://github.com/benmarwick/ctv-archaeology>).

line 810: Maybe also consider adding the diamond open access in the glossary as this concept is mentioned in the article.

Line 960: Within the question “How do I clean up the data and code before sharing this publicly”, I would suggest that the author can use one paragraph to first address the psychological barrier here that people are too ashamed to share their messy codes or afraid of potential criticism (<https://www.computerworld.com/article/2833340/4-reasons-developers-are-scared-of-making-their-code-public.html>). It is also important to emphasize that even researchers who have several years of programming experience will constantly seek help on Stack Overflow or similar platforms.

Line 1012: Language and package version should also be explicitly described. This is becoming increasingly important in R. The curse of the great ecology of R (many very specialized and ready-to-use packages) is that heavy package dependence of new packages makes them highly unstable. Maybe one tiny update of a dependent package will cause the dysfunction of the new package. Also, individual researchers who developed those small packages tend not to maintain them in the long term. For this reason, several researchers I knew started to use base R as much as possible or move to new languages like Julia. Perkel (2020) covered the irreproducibility of codes written years ago in a recent news piece.

line 1142: Another example of the diamond open access journal in archaeology would be PaleoAnthropology (<https://paleoanthropology.org/ojs/index.php/paleo/index>).

line 1225: Perhaps the authors can mention that there are specialized venues for publishing data like the Journal of Open Archaeology Data and Scientific Data. The former is designed for archaeologists, while the latter is a Nature portfolio journal that also accepts archaeological datasets like p3k14c (Bird et al. 2022) and SignBase (Dutkiewicz et al. 2020). As a side note, although impact factor is known as a highly problematic metric for research evaluation if exercised without caution, particularly within the open science framework, for those who do care about impact factor because of their university policy, Scientific Data actually has a higher impact factor (6.444) than Scientific Reports (4.379) or any archaeology journals. This number to some extent shows that publishing data is a behavior appreciated by the research community and can have a direct benefit to the authors.

Figure 1 and Figure S2: these two figures are not particularly informative.

#### **References:**

Bird, D., Miranda, L., Vander Linden, M., Robinson, E., Bocinsky, R. K., Nicholson, C., ... & Freeman, J. (2022). p3k14c, a synthetic global database of archaeological radiocarbon dates. *Scientific Data*, 9(1), 1-19.

Dutkiewicz, E., Russo, G., Lee, S., & Bentz, C. (2020). SignBase, a collection of geometric signs on mobile objects in the Paleolithic. *Scientific data*, 7(1), 1-14.

Gabelica, M., Bojčić, R., & Puljak, L. (2022). Many researchers were not compliant with their published data sharing statement: mixed-methods study. *Journal of Clinical Epidemiology*.

Perkel, J. M. (2020). Challenge to scientists: does your ten-year-old code still run?. *Nature*, 584(7822), 656-659.

**[Download the review](#)**