How can we measure and communicate the impact of science?

How can we measure the true impact of science? We're seeking feedback on indicators of the utility and rigor of publications beyond traditional journal metrics. Your input will help shape the future of our publishing experiment.

Version 2, published May 23, 2024. Originally published Mar 29, 2024.

Arcadia Science DOI: 10.57844/arcadia-74d6-huec

Purpose

Traditional signals of scientific quality — journal titles, closed peer review, and impact factors — don't fully reflect the utility and rigor of scientific work. Since our publishing platform exists outside of traditional systems, these signals wouldn't be available to us or those running other open science initiatives even if they were reliable. There are plenty of other challenges faced by scientists publishing both inside and outside of traditional systems too, including discoverability, tracking reuse, determining ways to re-evaluate quality over time when sharing living documents, and others.

We need new ways to evaluate science that better capture its true value *and* can be displayed directly on a scientific output so researchers can more easily utilize and expand on it.

The questions we've laid out at the bottom of this pub serve as conversation starters to creatively reimagine how we measure scientific efforts, especially forays into open science. We hope this dialogue will inspire us and others to develop open resources and tools that support science sharing for all collaborators in this space. Stay tuned for future publications where we'll share insights from our experiments with different reuse metrics.

Read on for background on what we've tried so far, or jump straight to the questions and start a dialogue.

- This pub is part of the **model creation effort**, "Reimagining scientific publishing." Visit the project narrative for more background and context on our approach to publishing.

Motivation

Research is most impactful when it's findable, accessible, and useful. Thus, a major goal of our publishing experiment is to release rigorous work that we and others can replicate and build upon. This is why we publish our science openly — complete with all the data, code, methods, and other information necessary to reuse and evaluate it.

Since we began iterating on our publishing framework [1], we've seen some early signs of success within and beyond Arcadia: community-driven GitHub contributions, reuse of our strains/reagents, alterations to preprints based on our modular reviews, and open feedback beginning to shape the way we think about our science.

Despite that, we are still working to identify all the indicators that will let us understand if we're meeting the goals of our publishing experiment.

What do we measure so far?

Strong metrics can inform our internal strategy and, when shared publicly, provide the people encountering our work with a means to quickly and effectively evaluate its usefulness. While we don't yet communicate any of this data to readers, we currently gather and analyze a variety of quantitative metrics, including:

Metrics about individual pubs

- PubPub:
    - Pageviews
    - Unique visitors
    - Country of visitors
    - PDF downloads
    - Number of public comments
    - Traffic sources
- Citations (via Google Scholar)

Metrics about linked resources

- Protocols.io metrics:
    - Views
    - Runs
    - Exports
    - Comments
- GitHub metrics:
    - Unique visitors
    - Unique clones
    - Number of pull requests (forthcoming)
    - Number of issues (forthcoming)
- Zenodo metrics:
    - Views
    - Downloads

We also gather qualitative metrics that could indicate utility and rigor, such as responses to the survey that you'll find at the bottom of every pub and public comments on our platform.

Tracking this data is helpful for researchers to determine who their work reaches, its quality, and how it's used. Still, it doesn't help readers understand if the work is rigorous or useful to them. We're developing ways to display metrics on our publications that reflect utility and rigor. But we're still figuring out the best form for that to take. If you have thoughts on what would be useful for you to see, please leave a comment here or on question number one!

What else do we want to measure?

While useful, many of the metrics above simply indicate reach (e.g. pageviews) or move at a pace that doesn't match ours (e.g. citations). Reach can be a useful marketing metric, but it doesn't reveal much about our science or its impact on its own. We need new ways to assess the utility of our work, ensure the feedback loop is fast enough to improve it, show scientific value to readers so they can quickly assess if a pub will be useful to them, and indicate how public feedback influenced our science.

What could we measure that would be more informative, and how would we collect that data efficiently? What parts of a pub is a given researcher using (code, protocols, data, etc.), and are they usable? How can we tell if our tools directly or indirectly inspire future work?

Many organizations and individuals are innovating in this realm; we aren't alone in this struggle. PLOS developed a set of "Open Science Indicators" to better understand the uptake of open science practices throughout the scientific ecosystem [2]. Recognizing the limitations of journal metrics, researchers in various fields have also proposed alternative frameworks. For example, the "Scientific Impact Framework" seeks to evaluate the influence of a piece of research using quantitative and qualitative metrics across multiple domains, from dissemination to implementation in public health policy [3].

And, with the rapidly expanding role of social media in facilitating scientific discussion, a variety of groups are working to gain new insights into who specific outputs are reaching and the dialogue surrounding them [4].

How might we continue to innovate together, share resources to document these efforts, and evaluate their outcomes?

Our goal is not to create a different impact factor — we recognize that scientific value cannot be boiled down to a single number and believe it should be conveyed through an array of different indicators. With rapid advances in AI and language processing, we as a science community are well-positioned to build nuanced, useful, and easy-to-parse methods to measure this.

Let's have a public conversation about how to identify and communicate qualitative and quantitative signs of rigor, utility, and reuse. We hope this forum will spark ideas for us and others to develop open tools or projects that will make it easier to evaluate scientific impact.

Weigh in!

While we'd love any thoughts or feedback you have, we've decided to focus on a small set of specific questions to provoke discussion:

1. In the absence of editorial decisions, what data, tags, summaries, or other information would help you quickly determine if a piece of research is relevant to your interests and use cases?
2. What existing or novel measures could indicate that research…
    - is verifiable (i.e., *can* someone verify that the work is rigorous and replicable)?
    - has been verified?
    - has been expanded or built on?

3. How might we effectively track the ways a given piece of research is reused (i.e., others following up on a finding, applying the knowledge provided, using a tool, etc.)? Are there existing tools that do this well?
4. What shared benchmarks should the open science community consider to evaluate the success of different publishing models?

If you like the idea of providing open feedback, consider weighing in on the questions above and signing up for our pub digest to get notified when we release new work! Remember, you don't need to write an entire review — we encourage in-line, modular feedback. Even a quick comment is appreciated!

Methods

We used ChatGPT to provide feedback on draft text and to suggest wording ideas and then used its responses as inspiration to improve the draft without directly using any of its phrasings.

Contributors (A–Z)

- **Prachee Avasthi**: Critical Feedback
- **Megan L. Hochstrasser**: Editing, Supervision
- **Jasmine Neal**: Writing
- **Robert Roth**: Conceptualization, Writing

References

1. Avasthi P, Hochstrasser ML. (2022). The experiment begins: Arcadia publishing 1.0. https://doi.org/10.57844/arcadia-050a-q254

2. Hrynaszkiewicz I, Kiermer V. (2022). PLOS Open Science Indicators principles and definitions. https://doi.org/10.6084/m9.figshare.21640889.v1

3. Ari MD, Iskander J, Araujo J, Casey C, Kools J, Chen B, Swain R, Kelly M, Popovic T. (2020). A science impact framework to measure impact beyond journal metrics. https://doi.org/10.1371/journal.pone.0244407

4. Carlson J, Harris K. (2020). Quantifying and contextualizing the impact of bioRxiv preprints through automated social media audience segmentation. https://doi.org/10.1371/journal.pbio.3000860