Published on Dec 19, 2024 by Arcadia Science **DOI**: 10.57844/arcadia-2a89-51c1

Early update on Arcadia publishing 2.0: Scientists are in charge, speed is an issue

Since starting v2 of our publishing model to restore scientist agency, pub “quality” has been similar, and this approach is more efficient overall. The major downside is that pubs take longer. We’re pursuing solutions to this and other problems but feel we’re on the right track.

Contributors (A-Z)

Prachee Avasthi, Audrey Bell, Keith Cheveralls, Megan L. Hochstrasser, Robert Roth, Wasim Sandhu, Kai Shanebeck

Version 6 · Sep 23, 2025

Purpose

About six months ago, we switched to a new “version 2.0” or “v2” model for our open publishing approach. We’d strayed a bit too far from our vision of researchers rapidly sharing their science on their own terms, and we needed to remove some top-down control over timelines, content, and process. Instead, our dedicated publishing staff (“pub team”) offers support on an opt-in basis.

This pub presents the initial results from this stage of our experiment. We've eliminated a lot of bureaucracy, and our scientists are now fully in charge. The pub team has more time to write meta-reflection pubs (like this!), improve publishing infrastructure and build tools, help more deeply on pubs that need it, and better integrate publishing with hiring. The main problem so far is that we're sharing less often. Without close oversight, not all of our scientists share promptly. We've also shifted a lot of responsibility to managers, who sometimes struggle to keep up with writing their individual pubs. Last, we haven't seen a change in the utility of engaging with the broader community — our scientists don't interact much externally on their own, and outside researchers seem to find and interact with our work a bit less often. Overall, though, we think these downsides are manageable, and we have plans to address them.

We based this pub on early data, after completing just 11 pubs via our v2 approach. We've since updated this piece to add a new section after releasing 18 more pubs. Most of our early indicators held true, and having more data hasn't changed our overall conclusions. The biggest difference is that, excitingly, our second wave of v2 pubs were significantly quicker. Taken as a whole, v2 pubs are still ever so slightly slower than v1, but if the current trend continues, it'll actually end up being faster.

We hope anyone following our publishing experiment will appreciate the update and give us feedback! We'd love to know what you think and hear any ideas for overcoming the issues we raise.

- This pub is part of the **model creation effort**, "Reimagining scientific publishing." Visit the model narrative for more background and context.
- The **data** we analyzed and the **code** we used to do so are available in this GitHub repository.
- Updated **data** and **code** from our September 2025 addition to this pub are available in this GitHub release.

Background on Arcadia's evolving publishing model

We've been publishing work on this platform, independently of journals, since mid-2022. Two years into our experiment, we reflected on an internal challenge that had slowly emerged — the deterioration of our scientists' agency in the publishing process — and decided to rectify it **[1]**. We refer to our original approach as "version 1.0" and our new strategy as "version 2.0" or "v2."

> We briefly summarize the key differences in our **new v2 framework** below, but check out our full pub on this evolution for more complete context.

In a nutshell, we realized that — though well-intentioned — our pub process had grown to be too complex and bureaucratic **[1]**. We required multiple meetings to plan and check in on pubs, and every draft had to go through review by the publishing team. In v2, we've shifted to a scientist-led system with just two requirements: byline contributor approval and availability of materials/information needed to reproduce the work. By default, our scientists now decide their own publishing formats, timelines, and engagement approaches without input from pub team. Our pub team still provides support services, but it's upon request from scientists, who choose from a menu of options.

We hoped this streamlined approach would better serve our goals of rapid, impactful scientific sharing while maintaining our commitment to openness and reproducibility. We figured the potential risks would be variation in quality, poor decision-making about commercialization by individual scientists, and lower pub output as a whole.

In this pub, we check how things are going with our v2 model. Read on to learn about the impacts we've seen so far!

The results

We're about six months in, and we've applied our new v2 publishing process to 19 total pubs (10 released, one we're keeping internal for now, and eight in progress). While we've been self-assessing progress constantly along the way, we thought now would

be a good time to take a more comprehensive snapshot of where things stand so we can reflect and update our readers.

We've mapped out our progress in Figure 1, manually approximating the relative size of each change we've observed to provide an overall sense of how things have evolved. In the subsequent subsections, we discuss what's going well and how we need to improve in more detail.

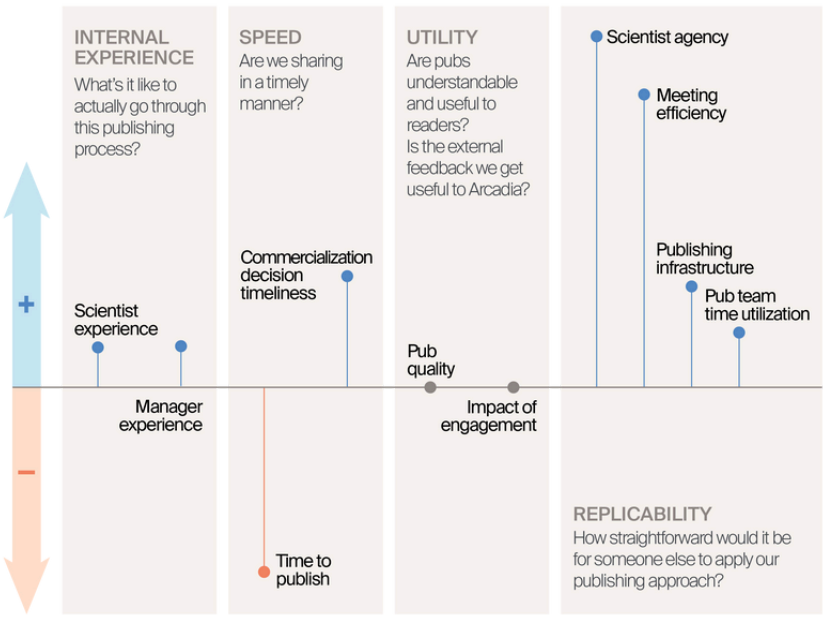

Figure 1

Summary of how things are going with our v2 publishing model compared to v1 as the baseline.

We've assigned rough ratings for how much various aspects of our publishing approach have changed in this new version of our model relative to our prior version. We group each readout based on how it connects to the big-picture goals of our publishing experiment — internal experience, speed, utility, and replicability. Keep in mind that these are approximations and only expressed relative to the original model.

What's going well

The process is streamlined and scientists are in charge

We've implemented a suite of changes that automatically streamline our publishing process and shift agency back to our scientists.

Things are much more efficient. We've eliminated two hours of required meetings for each pub and three hours of check-in meetings between the pub team and scientific team leads each month. We still do occasional ad-hoc meetings about particular pubs or sets of related pubs, but the number of these hasn't changed in our new approach. Unnecessary meetings are a huge opportunity cost for the company due to the people-time involved, so this reduction is a win for us.

Scientists now choose what help they get from the publishing team. Our standard v1 publishing process included about seven discrete tasks performed by members of the pub team for every pub (e.g., project management, editing the first draft, cleaning up figures to match our style, etc.). We've turned everything we used to do into an expanded menu of optional pub team "services." Each pub has unique needs. Our scientists can now choose what they want and skip what they don't. Thus far, each v2 pub receives around three services. One pub got seven services and another received none at all! Our scientists seem to have an intuitive understanding of what will help them most, and this can vary quite a bit from pub to pub (Table 1 shows our most-requested services).

Service	Description
Figure cleanup	Beautify figures, implementing style/branding guidelines
Thorough editing	Copyediting and in-depth feedback on content, structure, visuals, and overall effectiveness
Quick read-through for flow and clarity	Identify if any aspects are particularly confusing, suggest structural changes, provide feedback on visuals, and flag any key scientific details that seem to be missing
Original figures	Visualize methods, abstract concepts, organisms, processes, etc.
Technical consultation	Discuss what's possible on PubPub and beyond, explore technical feasibility of creative ideas
First-time pub onboarding	Meet-and-greet with pub team, overview of publishing at Arcadia and services offered, etc.
Visual strategy session	Talk through how visuals can support text (including data visualization, illustrations, graphical abstracts, and number/order of figures)

Table 1

Publishing services that our scientists have thus far requested at least twice.

Listed by number of requests, with most at the top.

In addition to our scientists being empowered, their perceived experience with publishing is also improved. After we release each pub, we send an optional survey to the researchers who were involved. Scientists who've published through both our original and new models report that their latest experience was the same or better. That said, we've only released a few pubs exclusively through the v2 system so far, and don't have many survey respondents (n = 6), so this is something we'll have to monitor over time.

Pub "quality" doesn't seem to have decreased

One concern about our new system was that with less oversight from our pub team, we'd start releasing pubs with more errors, confusing writing, or other issues. While "quality" is a difficult property to measure, we have a couple ways to track this, and

haven't noticed any major changes. We don't have a ton of responses, but we haven't seen a major change in how readers are responding to the form we include at the end of each pub asking about how clear, useful, and replicable it is. For example, based on the 214 responses we got for the pubs we released via our v1 process, each pub was rated "clear" about 85% of the time (as opposed to "a little confusing" or "confusing"). We have only 20 responses across our v2 pubs so far, but 83% of those responses called them "clear." Responses to other questions are also largely similar across v1 and v2 pubs. We also calculate a measure of readability (the "Flesch reading ease," or FRE, where higher numbers are easier to read — for reference, U.S. newspapers tend to have FREs around 45 to 50 **[2]**) for each pub using the scireadability Python package (v0.6.4) **[3]**. The average FRE is about the same for pubs developed via either our original or new model (~31 for v1 pubs; ~29 in v2). Both of these scores fall roughly within a US school level of "college" and "college graduate," respectively. While accurate syllable counting for FRE is challenging with heuristics, we believe these FREs are reasonably accurate and will continue encouraging our scientists to write in a clear, conversational style. We don't see a strong correlation between these data points and whether or not someone used pub team services. Again, we don't have a lot of data since we haven't released a ton of new pubs yet and "quality" is tough to define, but we haven't seen a huge qualitative difference either. We'll keep tracking this.

How have we maintained this consistent quality? Well, our scientists are pretty good communicators already. And in cases where a concept is tricky to articulate well or a writer is less experienced, they do a great job of getting the help they need from their manager, a collaborator, pub team, or a combo. With more free time, the pub team is even more available to help with pubs that need extra attention.

We've also implemented writing evaluations as part of our hiring process. We try to hire applicants who are both philosophically aligned with our open approach to sharing and able to demonstrate their ability to communicate science clearly. It's okay not to be a brilliant writer with perfect grammar, but the pub process is much easier for scientists who've mastered the fundamentals. We've also noticed that clarity of writing often correlates with clarity of thinking about complex topics, so it's a helpful evaluation point even outside of the publishing context.

Last, we started using Grammarly company-wide. In addition to ensuring basic grammatical correctness and improving the structure/flow of sentences, the business-level plan lets us create custom style rules to prevent common errors specific to science writing and to ensure stylistic consistency. This replaced a huge

proportion of the manual copyediting we used to do, and our scientists report other positive benefits, like making their writing more conversational, taking the second-guessing out of their process, and improving their writing quality.

We plan to experiment with AI-based writing solutions in the future, hoping not just to maintain quality, but to improve it further.

With less responsibility per pub, the pub team has more time to create core resources

Removing pub team tasks and meetings from our pub process means there's less of a built-in burden on our team's time. There are some clear, immediate benefits to our scientists — for example, the wait time for services is shorter, so the pub team bottlenecks pubs less often. We also have more time to invest in pubs that need close attention and help.

Having time freed up has let us reallocate it to bigger-picture pursuits that enhance our overall publishing experiment. Our publishing director has more time to help write "meta" pubs reflecting on our progress as a company (like this one!) and better incorporate writing evaluation into our hiring process. Our project manager has more time to quantitatively track the success of our experiment and build internal publishing infrastructure for our scientists (e.g., pub stats dashboards, automatic social media amplification of pubs, tools for finding relevant research, etc.). We've created guides on writing, accessibility, choosing the right pub type, checking first drafts for common errors, and more. We've also had time to implement new tools like Grammarly with extensive customized style rules and to collaboratively develop software packages to render figures in our house style **[4][5]**.

This makes our publishing model increasingly replicable. Now that it doesn't require a full support team, it's much easier to imagine a lab or a small startup trying our approach. Once we migrate to the upcoming new version of PubPub, called "PubPub Platform," we plan to share templates and many of the resources we've recently created to help others adapt our model.

Decision-making around commercialization has matured and hasn't delayed sharing

We previously had an "IP check" as part of our publishing process, and this arrangement occasionally caused delays in sharing content. While legal review of pub content only took a few days at most, punting decision-making about commercialization and patentability to the end of the scientific process could occasionally cause longer delays of several weeks because essential conversations were happening too late.

In our v2 model, we've removed the legal check and let our scientists decide what they share and when they share it. They're free to release pubs whenever they're ready, but also carry more individual responsibility to think critically about how we'll ultimately apply our research in the real world. This model helps encourage more forethought about downstream commercialization potential, during even the earliest phases of a new project.

Extricating commercialization-related strategy from publishing coincides with a similar shift elsewhere in the company. In a recent internal audit, our translation team found that rushing into pilots without planning around ideal outcomes or suitability for downstream therapeutic development was a common, preventable failure mode during their initial efforts to establish Arcadia's startup studio. Company-wide, we're now much more aligned on having scientists understand what it'll take to spin out a company and frontloading rigorous commercial research into their project plans.

Heartening, commercialization discussions haven't delayed the release of any v2 pubs. And it isn't just because we're sharing all of our work without thinking it through — we've completed a few pubs that our scientists expediently chose to keep internal until they finish the experiments necessary to understand whether their ideas have therapeutic potential. Our v2 shift and improved recordkeeping also led us to revisit a pub we decided to hold back during the v1 period, and we realized that we should make it public. So overall, the adjustments in our practices around commercialization seem to result in a default of openness, greater efficiency in sharing, and more thoughtful discussions on the key results we shouldn't share (yet).

One potential downside of this approach is that we'll accidentally release information that prevents us from patenting or commercializing a discovery. This doesn't seem to

have happened but will likely only become apparent in hindsight. We'll continue looking out for such occurrences.

Challenges

We have many ideas to address these problems, but it would extend the pub quite a bit to articulate them all. We'd love to hear any suggestions from you!

Pubs take longer, so we share less often

Let's address the biggest issue with our new system first. It's critical to our experiment that we share efficiently and openly at a pace that ensures our findings can benefit others quickly and that any feedback or ideas we receive are still actionable. The most worrisome trend in our v2 pubs is that they're taking significantly longer to complete. We previously took about 59 workdays to craft a pub from start to finish (mean, n = 55, standard deviation = 33). Pubs developed through our v2 process have taken 71 workdays (mean, n = 11, SD = 37). We also have eight draft pubs at the moment, and as of this pub's release, they've already been in progress for a mean time of 67 workdays. This is far too slow (for us, see below).

How long *should* a pub take?

It's important to note that our pubs take much less time to write than traditional research papers. It's hard to find reliable stats and this varies dramatically from author to author, but it can take anywhere from a few weeks to many months to write a typical manuscript **[6]**, let alone publish it. 71 workdays from typing the first word to seeing it live online is dramatically different.

For us, though, when a pub takes a "long time" to complete, it often means the point person isn't breaking their research down into modular "chunks" that they can share quickly, or it could indicate a scientific/performance issue. Delays also have major side effects: the process becomes inefficient with fits and starts that slow down other work, the contributors carry the nagging emotional burden of an unfinished task hanging over their heads, the external community misses out on timely information, and we can't make use of feedback if we've already moved onto other projects by the time we release and get critical comments.

We've seen how quickly a pub can come together when there's an external pressure, all without anyone working extra hours or feeling undue stress. So we don't want to let ourselves slide into a norm where pubs are unnecessarily slow. Based on our experiences so far, we'd say 20–30 workdays is fast for a data-centric pub, and 35–50 is a reasonable goal. Anything beyond 50 workdays is likely counterproductive.

Why are pubs taking so long? It's probably because the publishing team no longer provides automatic project management. Our team used to sit down with scientists and map out a plan and timeline for each pub. If someone fell behind, we'd remind them and guide them back on track. We'd also check in monthly with team leads to get a sense of all upcoming potential pubs and make sure they were kicked off in a timely manner. Without this close oversight, it's much easier for scientists to procrastinate, and the longer a pub starts to take, the tougher it usually is to finish.

We may actually be underestimating how long pubs are taking, too. For the pub team, the most challenging part of our v2 model is a lack of visibility. We can't tell when scientists have started writing a new pub unless they tell us or fill out our "pub kickoff" form. This happens most — but not all — of the time without prompting. We have almost no way of telling when new research could be shared, but the lead scientist, for

whatever reason, isn't writing it up. This limits our ability to track the success of our publishing experiment as a whole and makes intervention more difficult.

Many of the mechanisms we've developed for accountability depend on reliable tracking. For example, we have a "slow pub alert" set up for our CSO, a dashboard on a TV monitor in our headquarters that lists how long each pub has been in progress, etc. And though we no longer do automatic project management, we still provide occasional reminders. That said, none of these strategies are useful when we don't know someone's writing or *should* be writing a pub.

A few of our slowest pubs have lingered since before we switched to our v2 system. We count them as "v2 pubs" because they've been going through the new process, but in theory, these could have been written much earlier. Thus, some apparent overall slowness may stem from legacy pubs that suffer from problems independent of our publishing model. We hope that new pubs will move more expediently once we clear the backlog, but there are still changes we should make to speed things up.

Another factor contributing to this slow-down is that managers bear more responsibility in our v2 model, which can cause work overload.

Some managers are more overwhelmed

In our prior system, the pub team typically edited the first draft of each pub and its figures before other contributors reviewed the product. This could be slow, especially for less-polished pubs, because the pub team would spend a great deal of time puzzling through confusing information or seeking better ways to visualize data. In our new system, managers work with their reports to get first drafts into good shape before involving the pub team. Qualitatively, this appears to make the overall pub process more efficient because managers are most informed about what's not accurate or not well-described in early pub drafts and have the domain expertise to know immediately what to suggest instead.

The flip side of this efficiency gain is that managers have much more pub work. Strategizing about and reviewing pub drafts can create hours of work each week, especially for managers with larger teams or more junior team members. We also expect managers to execute and publish research independently, so they have to keep up with both their own pubs and those of their reports. Further, we've had a few instances in which managers must complete a pub for someone who's left the

company, which requires scouring over old notes and code of varying quality, completing analyses, and sometimes even redoing work.

Perhaps unsurprisingly, we've noticed that most pubs that aren't kicked off in a timely manner or that linger for months are those for which a manager is the point person. We don't yet have data from reports about whether waiting on their manager to help with their pubs tends to bottleneck their progress, but we'll track this moving forward.

That said, in a quick poll of affected managers, most find it no more difficult to manage their pubs than before. All rated v2 "better" or "much better" than v1 at both a conceptual level and in practice. They also report an overall better publishing experience in the new model, consistent with company-wide data.

The usefulness of external engagement is unchanged

A final major difference between our v1 and v2 publishing models is our decision to put our scientists in charge of engaging with other researchers about their work. Why engage? We want: 1) other scientists to know about our research so they can learn and build on it; 2) to get feedback from readers who catch critical errors or provide ideas, improving or speeding up our science; and 3) for readers to benefit from insightful comments from other readers. Rather than relying on pre-publication peer review, we've proposed that tracking how other scientists interact with our research will reliably indicate its usefulness and validity. To this end, the pub team previously coordinated engagement efforts for each pub, and at times, we tried having a team member dedicated entirely to facilitating interactions with the scientific community. In the v2 model, we've put this task entirely on our scientists. We hypothesized that we'd get more useful engagement with outside experts if we made it less like "homework," and our scientists felt the full responsibility and power of owning the process.

We've found that the impacts of our interactions with outside scientists are, in fact, unchanged. We send a handful of scientists to conferences every year, and our company recently held an "open house" event where we spoke more generally about our work with guests. Beyond that, scientist-driven engagement in our v2 system appears to be minimal.

While not the most reliable ways to understand how others benefit from our work, we've seen early signs that our v2 pubs receive fewer views on average and less

interaction (visits, clones, issues) with their associated GitHub repos. However, comparisons are difficult given the small number of pubs and GitHub repositories we've released in v2. We'll need to monitor this as we release more pubs.

Tracking public comments on our pubs also suggests infrequent or less effective outreach that doesn't lead to public feedback. We do receive new comments much more frequently than before, but it's because in August 2024, we started requiring anyone applying to a scientific role to leave a comment as part of their evaluation. Before this, the median number of comments we received each month was four — now it's 25! Almost all pub comments (~90%) now come from job applicants rather than readers who discovered our work organically or were made aware of it via outreach from our scientists.

Are the comments more or less helpful now? We have scientists rate the impact of comments they receive on their research (choosing from one or more options like "No impact," "Changed our direction/idea we're following up on," and "Caught typo/small error"). Though we have fewer rated comments for our handful of v2 pubs so far, the resulting impact profiles are roughly consistent between v1 (n = 223 rated comments, 231 comment impacts) and v2 pubs (n = 28 rated comments, 29 comment impacts) (Figure 2).

In summary, with less outreach to external scientists, fewer people see and build on our work. We do get more comments because we require this of job applicants, but they're not incredibly useful outside of the hiring process.

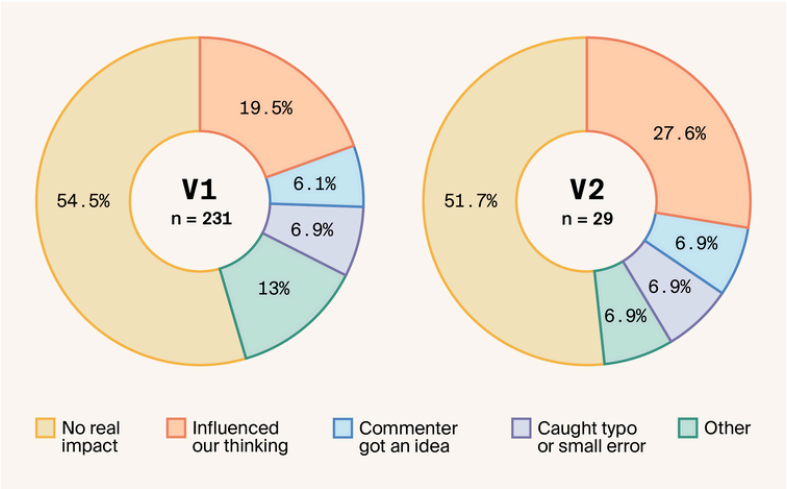

Figure 2

Comment impact breakdown by publishing version.

When our scientists assign impact categories to public comments, they can choose multiple categories per comment if needed. Here, we split comments tagged with multiple categories, meaning we count each individual “impact” and calculate percentages based on the total number of impacts rather than total number of comments. “Commenter got an idea” denotes a perceived benefit to the commenter themself rather than to our scientists. “Other” is any impact that appeared fewer than 10 times across v1 and v2 categorizations.

The similarity between these distributions is an early suggestion that our shift to scientist-led engagement hasn't meaningfully changed how public feedback shapes our work.

Methods

> The **data** we analyzed and the **code** we used to do so are available on GitHub (DOI: 10.5281/zenodo.14969194).

We calculate workdays in progress by recording the “kickoff date” of the pub (when we held the kickoff meeting for v1 pubs and when the kickoff form was submitted in v2) and the publication or internal completion date. We use the `WORKDAY_DIFF()` function in Airtable to calculate an approximate number of workdays in progress. This doesn't account for individual PTO, company events, or other external factors beyond weekends. We used this script to calculate summary statistics related to pub duration.

To generate the donut charts and associated data in Figure 2, we used data stored in Airtable and this script (modified here to accept a CSV instead of using the Airtable API) to calculate the number of times our scientists assigned a given impact. We also used the Airtable summary bar to calculate how many comments were submitted by job applicants after our switch to this requirement. The pub’s point person assigned or approved all comment impacts.

We calculated a Flesch reading-ease score for each pub using scireadability (v0.6.4) in this script.

We used Grammarly Business to help clarify and streamline our text. We also used Claude (3.5 Sonnet) to suggest wording ideas and then choose which small phrases or sentence structure ideas to use, to write longer stretches of text that we edited, and to help write and debug code. We used Gemini 2.5 Pro to help write and clean up code for our September update to this pub. We also used GitHub Copilot to help clean up code.

Key takeaways

The good

- The process is streamlined and scientists are in charge

- Pub “quality” doesn’t seem to have decreased
- With less responsibility per pub, the pub team has more time to create core resources
- Decision-making around commercialization has matured and hasn’t delayed sharing

The bad

- The usefulness of external engagement is unchanged
- Some managers are more overwhelmed

The ugly

- Pubs take longer, so we share less often

Fast-forward: How have things changed more than a year into v2.0?

We released the first version of this pub in December 2024. We’ve added this new section (and made some tweaks to the “Purpose” and “What’s next”) eight months later, after collecting more extensive data on our publishing v2 model. This update covers pubs released between December 15, 2024 and August 15, 2025. In that time, we released 18 new pubs, bringing the v2 total to 34. While we didn’t repeat all our initial analyses, we walk through some more fully informed conclusions in the following subsections.

Pubs are starting to go much faster

TL;DR

There isn't a significant difference in speed between v1 and v2 pubs. But v2 pubs appear to be getting faster as time goes on.

Overall, the average time to publish for all published v2 pubs is 60.8 workdays (n = 34, SD = 42.6, 95% confidence interval [45.9, 75.7]). This is slightly slower than the v1 average of 56.4 workdays (n = 54, SD = 32.4, 95% CI [47.5, 65.2]), but the difference isn't significant (p = 0.606, Welch t-test).

We're encouraged by what we find when we look at the "initial" cohort of v2 pubs and the ones released since our original update. The initial 16 v2 pubs we originally reported on (some of which were in progress at the time and have since been published) took an average of 84.4 workdays to complete (SD = 41.4, 95% CI [62.4, 106.5]). The 18 "new" v2 pubs (kicked off and published since our last update) were completed in an average of 39.8 workdays (SD = 31.9, 95% CI [23.9, 55.7]). This represents a statistically significant improvement in publishing speed of about 44 workdays (p = 0.0016, Welch t-test). We're still dealing with small sample sizes, but this is a very positive direction.

We believe there are a few reasons for this change. One is that we began publishing "notebook pubs" — findings shared directly as computational notebooks rather than sharing code and narrative split into separate artifacts **[7]**. While we haven't released enough of them to say if they're significantly faster as a rule, many of our quickest pubs are now notebook pubs.

Also, people are getting used to v2, and our tooling has evolved. We've chipped away at certain required steps, increased AI and automation throughout the process, and more of our scientists have published in v2 and gotten comfortable with the tools we have available.

These comparisons between the speed of v1 and v2 should be made with caution, as we need to grapple with a number of confounding variables that could increase or decrease the amount of time a pub takes — personal factors such as comfort with writing, company pivots, the requirements of a pub and its associated research

artifacts, personnel changes, scheduling difficulties, and so many more. We don't interpret these trends as solely being related to the difference between v1 and v2, but they're useful to track as we make decisions about our publishing experiment.

We'd also like to thank Kai Shanebeck for their comment on this pub regarding the confidence intervals for pub speed. We've added those stats here to help further clarify the speed differences between v1 and v2.

Quality and external engagement metrics are similar

> **TL;DR**
>
> Scientist-driven engagement remains low in v2, and much of our direct engagement on pubs comes from job applicants. Imperfect "quality" measures are mostly unchanged, but v2 pubs show a different distribution in feedback form responses.

In our first update to v2, we reported on two metrics that can provide insight into the "quality" of pubs: Flesch reading ease (FRE) and responses to our feedback form at the bottom of each pub. The average FRE of v2 pubs (27.6; n = 34) is still roughly the same as v1 pubs (30.8; n = 54), and this difference isn't statistically significant (p = 0.072, Welch t-test). However, we do see some shifts in the feedback forms we receive.

We have a larger sample of responses now (n = 444; 267 for v1 pubs and 177 for v2 pubs). Note that all questions are optional, so n can differ between questions. We also treat responses as independent in these tests and interpret these as exploratory analyses, not definitive conclusions. We don't see a significant change in how readers rate the ability to find everything needed to reuse/reproduce the work (n = 429, χ^2 test, p = 0.284) or the evidence supporting claims (n = 356, χ^2 test, p = 0.181). There are other shifts, though.

First, the overall "clarity" rating remains high for both v1 and v2, but the distribution of unclear responses has shifted somewhat. Readers of v2 pubs were less likely to label them as "Confusing" (3% vs. 6.6% for v1) and more likely to select "A little confusing"

(13.2% vs. 6.6%) (n = 423, χ^2 test, p = 0.025). That said, this isn't significant after multiple testing correction (Bonferroni p = 0.101).

Second, the overall distribution differs in the “usefulness” rating of v2 pubs. Proportionally, “Maybe” useful responses are higher (41% vs. 29.0% for v1) and “Not really” useful responses are lower (11.6% vs. 21.2%) (n = 432, χ^2 test, Bonferroni-adjusted p = 0.024).

Our main metric of engagement, the number of reader comments, is still higher in v2 than v1 and about 90% still come from job applicants. Other indicators of engagement, like GitHub repository visits and clones, are increasing. So are general pageviews, comments, and citations. In the first six months of v2, it appeared that there was some slowdown in GitHub visits and clones, but we now observe that v2 repositories are generally receiving a similar or higher number of clones on a per-week basis, but visits are more difficult to pin down. Rather than relating to any differences in our v1 vs. v2 processes, we believe these differences are more likely related to an increase in traffic generally, greater awareness of our pubs, our requirement that job applicants leave a comment, and varying interest in our different GitHub repositories based on the nature of the content.

A note on comments from job applicants

While our requirement for applicants to comment on our pubs is somewhat unconventional, we've found it can generate high-quality engagement. Though many comments are indeed perfunctory and we do select for overly positive and uncritical responses, some include substantive critiques, detailed reanalyses, and thoughtful extensions of our work that provide genuine scientific value. This approach lets us leverage an existing process to sometimes capture useful feedback without additional effort, and we think it could be even more effective when supplemented with targeted outreach for a fuller picture of community response.

Commercialization decisions still aren't delaying pubs

> **TL;DR**
>
> No v2 pubs have been delayed by commercialization decisions. We've also released all the v1 pubs we'd previously held back to consider for commercialization.

As in our original update, no v2 pubs have been delayed by commercialization decisions. While multiple pubs were "paused" by a scientist after kickoff, none of these delays have been related to commercialization. Additionally, all the pubs that we were holding internally have since been released publicly after discussions with the scientists who worked on them, supporting our hypothesis that overly cautious approaches can delay valuable community feedback without ultimately protecting IP. In a few cases, we found that earlier publication would have been beneficial not only for external feedback but also because the process of preparing work for publication itself routinely helps identify issues and strengthen research. Though changes in company direction have also contributed to this, we view increased clarity around IP as a continuing positive sign for publishing v2 and will continue monitoring.

A note on data

You may notice discrepancies between the numbers we report in this update and our original publication. We've since improved our methodology for counting workdays for paused pubs, which caused small shifts in workday counts. We've also now released all the pubs we were holding internally — some of these pubs, which we counted as "complete" pubs in our first report, took additional time when authors revisited the content prior to release. Additionally, we decided not to publish one "internally complete" v1 pub from the initial dataset, but instead wrote a new pub with a different focus based on the same data using the v2 process. Further, we omitted two pubs that were included in our original analyses because they were kicked off during v1, work stalled, and then they were restarted and released during v2, thus going through both publishing processes.

All **data** for our **September 2025 update** (and a few additional analyses) are available in this folder of our GitHub repo, and all **code** is here.

All data from our **original release** of this pub are in this folder, and all **code** is here.

Are the trade-offs worth it for us?

In a word, yes. We have a lot of ideas to improve, which we've omitted for space and because many are pretty specific to Arcadia's inner workings. But even if all our attempts to improve fail completely, the benefits still outweigh the downsides (Figure 1), and we're sure this new strategy is a move in the right direction. Scientist agency and model replicability are critical to the success of our experiment, and we'd strayed too far away from those guiding pillars. We may eventually devise yet another new model for publishing, but we won't go back to our original approach.

That said, we're hopeful that by implementing our improvement ideas, and perhaps by gathering constructive insights from you, our readers, we can get even more creative.

How might *you* benefit from what we've learned?

We hope you've found this pub interesting, but ultimately want you to take something useful away from these reflections. Even if you're not exploring new models of research sharing, our experiment has surfaced some practical insights that might inform your publishing practices. Our high-level lessons are probably most useful for group leaders or people taking the lead on prepping a manuscript.

Balance agency with accountability

While giving authors full control over publishing can foster creativity, some structure helps maintain momentum:

- Consider providing more formal training in writing rather than relying purely on an apprenticeship model where junior scientists learn via observation, or, when giving feedback to authors who are learning, make sure to include comments specifically on how to improve their writing
- Light accountability or project management may be the best compromise between heavy involvement and a hands-off approach — consider regular asynchronous check-ins and occasional in-person meetings with new writers, like grad students
- Establish clear, realistic timeframes for text and figure drafts, as well as overall completion of the manuscript
- Ask the lead author(s) what support they think they'll need most before you start and check in routinely to see how this changes
- Consider tracking time spent on manuscript preparation to identify and address bottlenecks, or at least reflect on this often
- Set up lightweight practices to flag when projects are moving too slowly or signaling problems unrelated to writing
- Start writing sooner and include input from managers/PIs earlier to prevent a heavy lift at the end

Lean on tooling and automation

By embracing the right technological solutions, you can reduce manual workload and free up you/your team's energy for higher-value work:

- Tools like Grammarly can help a lot — it could be worth paying for Grammarly Business and creating custom style rules for your lab or group to cut down on time spent fixing basic issues
- Automate as much as possible! Pre-schedule reminder emails, use tools like Geekbot on Slack to get weekly updates on writing progress, create templates or

scripts for quick figure design, automatically re-up social media posts about papers, etc.

If you want feedback, engage

If you're hoping the scientific community will help improve your work through discussion, there are a few ways to focus their attention:

- While extensive outreach strategies don't always lead to extensive feedback, silence essentially guarantees zero engagement. Quick social media posts, emails to colleagues, requesting a PREreview or feedback from us on your preprint, and presenting at supergroups or local symposia are all relatively low-lift ways to make people aware of your research and to encourage critical discussion.
- Ask applicants to your lab or company to leave public comments on your preprints or papers. Some sites have built-in comment functionality (e.g., Disqus on bioRxiv), Hypothes.is lets users comment on any URL, and you might even suggest more extensive feedback mechanisms like PREreview or preLights. This isn't just helpful in polishing your science; it also provides deep insight into a candidate's mindset, expertise, and suitability for a role. Plus, it gives the candidate a chance to consider more thoughtfully whether this is the type of work they'd like to pursue next.
- Give someone else feedback and then ask for it in return. If you need inspiration, we curate a list of preprints for which authors have asked for public feedback. Once you've added your thoughts, let the corresponding author know and suggest a piece of your own that could benefit from similar open review! Even if they don't reciprocate, by publicly commenting on someone else's science, you've provided a valuable service to both the authors and other readers.

What's next?

We'll keep tracking our publishing experiment and update you when we have new findings, ideas, recommendations, or need input. We'll also let you know when we step into our next iteration, be it incremental or perhaps a full v3.

We'd love to hear what you think about our initial observations and ideas to improve our v2 publishing model. Are there other aspects of this strategy that we should measure to understand our overall success? What should we do to address the pub slowdown and lack of external engagement?

References

1. Avasthi P, Hochstrasser ML, Roth R. (2024). The experiment continues: Arcadia publishing 2.0. https://doi.org/10.57844/ARCADIA-1C73-8CB5

2. Flaounas I, Ali O, Lansdall-Welfare T, De Bie T, Mosdell N, Lewis J, Cristianini N. (2013). RESEARCH METHODS IN THE AGE OF DIGITAL JOURNALISM. https://doi.org/10.1080/21670811.2012.714928

3. Roth R. (2025). scireadability. https://github.com/robert-roth/scireadability

4. Arcadia Science. (2024). arcadia-pycolor. https://github.com/Arcadia-Science/arcadia-pycolor

5. Arcadia Science. (2024). arcadiathemeR. https://github.com/Arcadia-Science/arcadiathemeR

6. How to write a research paper. (2023). https://doi.org/10.1126/science.caredit.adi0662

7. Avasthi P, Bigge BM, Hochstrasser ML, Kiefl E, Roth R, Sabbagh U, York R. (2025). Closing the divide between analysis and publication: The notebook pub. https://doi.org/10.57844/ARCADIA-CA21-23BB
