Open Practices in Visualization Research

Opinion Paper

Steve Haroz
oavis@steveharoz.com
Sorbonne Université

ABSTRACT

Two fundamental tenants of scientific research are that it can be scrutinized and built-upon. Both require that the collected data and supporting materials be shared, so others can examine, reuse, and extend them. Assessing the accessibility of these components and the paper itself can serve as a proxy for the reliability, replicability, and applicability of a field's research. In this paper, I describe the current state of openness in visualization research and provide suggestions for authors, reviewers, and editors to improve open practices in the field.

A free copy of this paper, the collected data, and the source code are available at https://osf.io/qf9na/

1 THE CURRENT STATE

In 2017, Open Access VIS (http://oavis.steveharoz.com), surveyed what information is openly accessible for all papers presented at the VIS conference. The goals of the site were to:

1. Collect openly accessible papers, data, and materials from VIS papers before the conference.
2. Encourage authors to post papers, data, and materials by highlighting those that do.
3. Encourage posting on reliable open access repositories to avoid "link rot".
4. Get an overview of how many VIS papers are inaccessible outside of a pay wall.

1.1 Information collected

Besides publication and conference metadata, Open Access VIS (OAVIS) focused on primary research components needed to replicate, reanalyze, or apply the results in the paper.

An open access copy of the paper

The VIS publisher, IEEE, allows authors to share the post-reviewed version of a published article on their personal and institutional websites and on preprint servers. Papers posted by non-authors or on sites without clear licensing (e.g., researchgate.net, sci-hub.tw, or semanticscholar.org) were not included. The ability to read the paper is critical to understanding the context of other links and artifacts (such as data or materials), so the additional content was only shown if a free legitimate copy of the paper was shared.

Materials

While authors often share a variety additional media, this badge was only awarded if the material shared could allow an independent person to completely replicate critical components of the results. Example components that satisfy the criteria include experiment source code, experiment stimuli, a questionnaire, or a framework's source code. Any datasets needed to recreate figures or experiment stimuli should also be considered "materials". This category is an implementation of the Center for Open Science's Open Materials badge, whose full criteria are available at https://osf.io/gc2g8/

Data

Data is an ambiguous term for visualization research. In the context of open science, "open data" refers to the raw data measured or categorized by the authors to support the conclusions of an empirical study. However, visualization papers often use datasets to generate experiment stimuli or demonstrate example applications, in which case, the dataset is a component of the *materials*, necessary for replicating an experiment or reproducing figures. In the open science context, data can be thought of as "raw results". Examples include the per-trial results of a human-subjects experiment or timing and memory consumption measurements from an algorithm comparison experiment. The full criteria for the Center for Open Science's Open Data badge are available at https://osf.io/g6u5k/

Explanation

Many VIS papers are also accompanied by a site that explains or demonstrates the contribution to a more general audience. It can often act as an author-written press release for the project. Compared to the other open practices catalogued by OAVIS, this practice is an outlier, as it is not part of the research and therefor not critical for replication. However, given that many groups create these sites and demos, understanding their makeup and reliability is potentially useful, especially since people have independently collected data on these pages in the past (http://www.aviz.fr/~bbach/publicize_vis_2016/). I included a criterion that an explanation site should provide information beyond the basic metadata of the paper and resources that are catalogued by OAVIS. Therefore, websites were included in this category only if they went beyond a simple landing page and provided some explanation or a demo.

1.2 Long-term reliability

An important requirement for the Center for Open Science's Open Data and Open Materials badges is that the materials "*are publicly available on an open-access repository, [and they] must have a persistent identifier and be provided in a format that is time-stamped, immutable, and permanent (e.g., university repository, a registration on the Open Science Framework, or an independent repository at www.re3data.org).*" This requirement exists because personal websites, institutional websites, and personal repositories are rarely persistent or reliable over a long period of time [1], and changes are not transparent. Students graduate; sites are reorganized; or content gets deleted. This "link rot" can be very problematic for anyone trying to apply, extend, or replicate the research years after publication. A potential source of future link rot is seemingly stable commercial services that have no plan for long term persistence. The Open Science Framework (OSF - osf.io) is one service that has a persistence plan available at http://help.osf.io/m/faqs/l/726460-faqs#what-if-you-run-out-of-funding-what-happens-to-my-data

Some universities libraries have their own open-access repository. However, they often suffer from limited discoverability, and there is not a reliable way to determine if a given university has a repository or if a given URL is on a repository or some professor's

website. SHARE is a project that is attempting to provide a unified portal both to the larger and well-known repositories and preprint servers as well as smaller ones hosted by individual universities - `https://share.osf.io/`

Very few VIS papers make use of reliable repositories. To incentivize at least some sharing of the paper, data, and materials, a "partial credit" version of each badge was awarded to if the information was shared on a less reliable site. Unfortunately, 5% of these unreliable paper links suffered from "link rot" within two months. Those papers were removed from OAVIS and excluded from this analysis.

1.3 Information not collected

A preregistration is a plan for an experiment that is timestamped prior to collecting data. A cursory review of the available papers did not reveal any preregistration. However, they can be difficult to search for, so it is probably infeasible to collect preregistrations unless authors disclose it in the abstract or the submission form. An example disclosure form for Center for Open Science badges is at `https://osf.io/5fndw/`

Supplemental material was not linked to in OAVIS because the content and its value are often unclear. It could be as trivial as a copy of all the figures. Also IEEE keeps it behind a paywall with an unclear license. In a cursory review, the clearly labeled artifacts defined by the open science badges appear to largely cover most of what authors include in supplemental material.

1.4 Openness overview

143 papers were presented at the conference in 2017. Of those, only 21 (15%) were available on a reliable open-access platform. However, the majority (71%) were available in some way, including on an author or institution website (Figure 1). Sadly, 43 papers were not shared even though three months passed from when the final versions were sent to the publisher to the beginning of the conference. Some of the unavailable papers were accessible but only on sites that do not appear comply with the publisher's license.

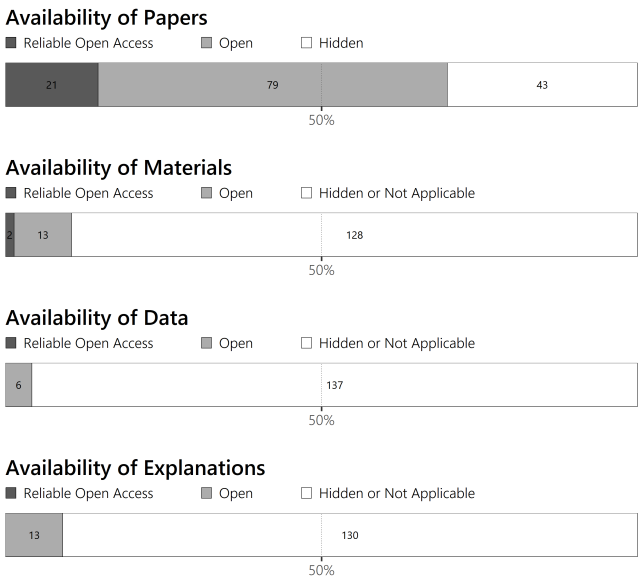

Figure 1: An overview of availability of VIS papers, data, and materials.

The often touted applicability of VIS papers makes the lack of shared materials troublesome. The primary contribution of many papers is a library, technique, algorithm, or application, so if these resources are not shared, it is unclear how people in academia or industry can make use of the researchers' efforts.

The low number of papers with open data is also concerning. Although many papers do not include any experiment or measurement, the InfoVis track alone includes 15 papers with experiments with objective measures and even more papers with other types of measurements. Yet only 6 papers shared raw data (none reliably).

Furthermore, the vast majority of artifacts were posted to an unreliable location without any guarantee of persistence. It is likely in that in the coming years, many that are hosted on personal or institutional websites will disappear. For materials especially, many repositories are hosted on GitHub, which could restrict access at any point (it is unlikely, but not impossible). So the future availability is perilous.

1.5 Openness by review venue

Figure 2 shows that across the three main tracks at the VIS conference (VAST, InfoVis, and SciVis), none had more that 15% of papers shared on a reliable repository. However, looking at paper sharing irrespective of reliability, the InfoVis track had the largest sharing rate of 85%, compared to VAST (65%) and SciVis (52%). Further break down of the three main tracks shows that more InfoVis papers shared materials (6), data (4), and explanation pages (6) than VAST and SciVis papers combined. The small sample size and different publication process makes the journal-reviewed articles (TVCG and CG&A) difficult to compare.

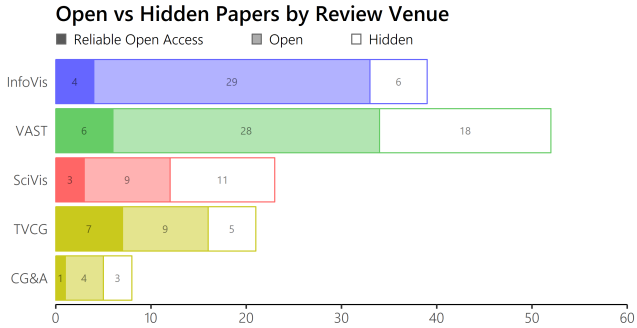

Figure 2: Different review tracks and the cultures underlying them yield fairly different rates of paper sharing.

2 How to Increase Open Practices

Research is mostly publicly funded, performed by students, and reviewed by volunteers, and hiding the fruits of that labor (publications) or putting them behind a paywall only hinders the efficacy of research progress and prevents the field's ability to self-correct [2]. Failing to make research data and material available also requires the research community to blindly trust that the authors made no mistakes, were fully transparent, and never acted dishonestly. Ultimately, it is not clear why researchers should have to buy back the publication they wrote and reviewed for free while blindly trusting in the perfect behavior of their colleagues.

The arguments for openness are not only philosophical and egalitarian but also practical for the authors. Allowing others to freely access a paper [3] and research data [4] is correlated with increased citations. Everyone in a research community can contribute to improving the openness of the field's research.

2.1 Advocate openness to counter inertia and apathy

The ability to share materials is a relatively recent phenomenon on the timescale of research practices. For more senior researchers, the ability did not exist when they were students, so without any suggestion to open their research, they're riding the momentum of

an antiquated approach to research. Yet digital accessibility can improve the efficiency of dissemination and the veracity of published research [2]. Vocally advocating for open practices may help inform and convince more people.

OAVIS provides some anecdotal evidence that advocacy works. Although only two papers' authors posted material to a reliable open-access repository (Figure 1), both cited OAVIS as a reason. Providing more clear benefits and promotion for research posted to a reliable repository (perhaps by limiting the OAVIS features for papers with less reliable links) may help improve open practices.

2.2 Incentivize with Open Science Badges

The Center for Open Science created badges to show the openness of multiple research components (https://cos.io/our-services/open-science-badges/) [5]. These badges are rewarded for three open practices:

- Open Materials: Earned by sharing research materials needed to replicate the results. Full criteria: https://osf.io/gc2g8/
- Open Data: Earned by sharing all raw data measurements and analysis scripts from an experiment necessary for reproduction of analyses. Full criteria: https://osf.io/g6u5k/
- Preregistered: Earned by sharing an experiment plan that is timestamped prior to collecting data to minimize questionable research practices [6]. Full criteria: https://osf.io/7d4wa/

The badges have been implemented by dozens of journals (https://osf.io/tvyxz/wiki/5.%20Adoptions%20and%20Endorsements/). Early evidence shows that adopting the badges substantially increases open practices among papers [7], as a journal that implemented the open data badge saw substantial increase in data sharing compared with other journals in the same field. The Center for Open Science has resources and template disclosure forms for authors to help publication venues implement the badges at https://osf.io/tvyxz/

2.3 Conference and journal policy

The Transparency and Openness Promotion (TOP) guidelines comprise eight categories of conference/journal or funding policies to promote open practices in published papers. The categories are:

- **Citation** standards for data, code, and other materials to ensure that contributions outside of a paper are respected and cited appropriately.
- **Data** guidelines regarding raw research results.
- **Analytics** guidelines regarding reproducible analyses.
- **Materials** guidelines regarding materials needed for replication.
- **Design and analysis** guidelines for thorough reporting of procedures and analyses. Guidelines and exemplars for statistical reporting in HCI [8] and templates for experiment methods reporting [9] could serve as starting points.
- **Preregistration of studies** for reporting any prespecification of experiments prior to data collection
- **Preregistration of analyses** for reporting any prespecification of analyses prior to data collection
- **Replication** for the venue's policy towards and handling of replication submissions and ensuring that publications are replicable

Each category has three levels of strictness, and publication venues can choose the level of each category independently. The levels generally correspond to:

1. Disclosure of whether and where materials are available
2. Require open materials or an explanation for their absence, but trust the authors to affirm
3. A process for review of submitted materials

This flexibility gives publication venues a low barrier of entry and the ability to progress gradually. For example, the journal *Science* adopted level 1 (transparency of whether it was done) for preregistration and replication, but it adopted level 2 (mandatory submission or a stated reason for not submitting) for data, analytics, materials, and design and analysis [10].

Critically, the guidelines allow for exceptions. For example, if sharing the raw data would violate privacy or intellectual property rights, explaining that reason the paper would satisfy the TOP requirements. Transparency is key, so the location of an external resource or an explanation for why it cannot be shared must accompany the paper.

As of June 2018, over 700 publication venues from a broad range of disciplines have adopted the TOP guidelines. The journal, Cognition, found that standards mandating data submission (level 2) resulted in a measurable increase in the reproducibility of analyses [11].

The website and repository for the TOP guidelines include resources to aid in adoption:

- A summary of each level for each category – See Appendix A
- The full text of each level for each category – https://osf.io/9f6gx/wiki/Guidelines/
- Sample implementations for journals and funding agencies – https://osf.io/9f6gx/wiki/Sample%20Implementations/

2.4 The Peer Reviewer Openness (PRO) initiative

Reviewers can help promote open practices in the venues they review for. Many reviewers have encountered questionable practices in submissions they review but did not have sufficient implementation or data details to adequately examine their concern. While no one should be obligated to trust authors, that is especially true for reviewers whose primary obligation is to scrutinize claims.

The Peer Reviewers' Openness (PRO) initiative [12] is pledge for reviewers to take matters in their own hands and progress openness and transparency irrespective of the publication venue's policies. Reviewers accept review invitations as they normally would, but only perform a full review if minimum requirements for open data and materials are met. Otherwise the submission is reject without being read. The requirements are similar to the level 2 TOP requirements for data and materials and are available at https://opennessinitiative.org/the-initiative/

Like the TOP guidelines and open science badges, reasonable exceptions are allowed for. If data and materials cannot be made available (e.g., privacy concerns or because some of the data is being used for in-progress research), the reason must simply be stated in the submission. Even if a submission fails to follow the requirements, the initiative instructs reviewers not to reject immediately. Instead, it provides a straight forward script to ask the editor (or primary reviewer) to send the authors a request to meet the requirements along with helpful advice for doing so. The goal of the initiative is to improve the quality of research and to be as helpful as possible.

2.5 A necessity for career advancement

Funding agencies and research institutions can be effective in advancing openness in research [13]. Funding agencies can stipulate that all research funded through its grants are published in open-access venues with open data and materials. There is a version of the TOP guidelines specifically for funders. Research institutions can also promote open practices among its staff by educating them on open science and promoting it. Some universities have even begun requiring that researchers demonstrate the openness of their research before being hired (http://www.nicebread.de/open-science-hiring-policy-lmu/). More generally, people who serve on hiring or tenure panels can make a habit of reading applicants' work from outside a

university paywall and checking if the paper, materials, and data are accessible.

3 How authors can include data with anonymous submissions

While there are a variety of guides for sharing data and materials openly and publicly (e.g., [14]), the approach is not entirely compatible with the anonymous submission approach that is commonly taken in VIS. The Open Science Foundation (OSF) supports private repositories with a special URL that strips out names to be anonymous to reviewers.

3.1) Make an Open Science Framework account.
`http://osf.io`

3.2) Make a new project. Give it a name, add collaborators, and fill in other details. You can update everything later, too.

3.3) Upload your content. Encapsulate each experiment's materials, data, and (if possible) analysis in its own subdirectory. Also, put a current draft of the paper at the root.

Post all the materials for each experiment, including the code, parameters, and any special running instructions. Think about what someone 5, 10, or 50 years from now would need to replicate your work.

For experiment data, make sure to also post a dictionary that describes every field. Posting your analysis script along with the data is a helpful addition that can improve replicability and help justify your results.

3.4) If anonymity is a concern, make an anonymous view-only link. With your project open, go to Settings (top-right) ➔ View-Only Links ➔ Add. Then check "Anonymize". Then you can share that URL. See Figure 3.

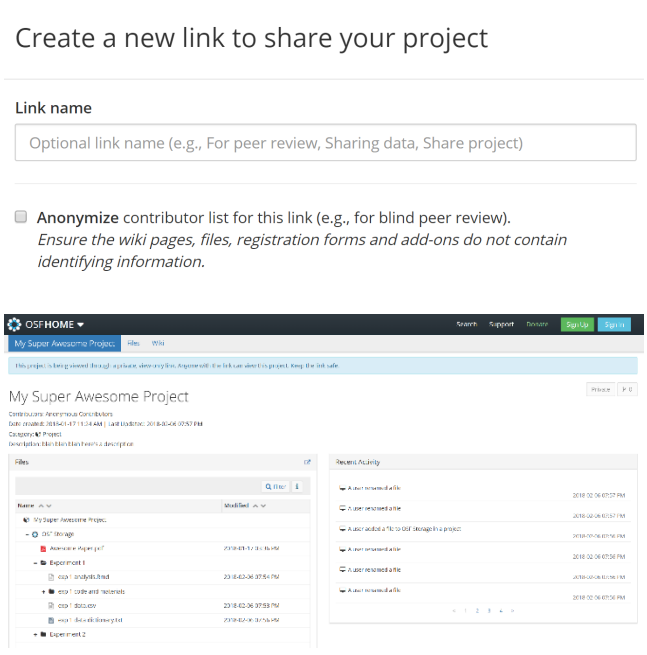

Figure 3: (Top) A screenshot of the option to generate a link to an OSF project and anonymize all names in the project metadata. (Bottom) A screenshot of an OSF project viewed via an anonymous link. Note: the anonymization only applies to metadata, not file content.

3.5) Add the link to your abstract. Don't bury it deep in the paper or all the way at the end. Put it on the first page! Putting it in the abstract rather than the body can allow people to access everything even if the paper is behind a paywall. Note that the URL is much shorter once you make the project public, so don't worry about the length.

3.6) Make it public. After the submission is accepted (or before!), make the project public. That will give you a shorter URL and will reveal all the collaborator names.

4 How authors can share research results effectively

Even with good intentions, many missteps can happen that reduce the value of open data. These tips should help ensure that research data is indeed open, useful, and accessible.

4.1 Provide at least the minimum information

Experiment data should include an entry for each trial, usually as a row in the table [15]. It should not be aggregated by condition or by subject. Here is a minimum set of columns that would be appropriate for most experiments:

- **Subject ID** – Remember to keep it anonymous. No names, mTurk IDs, or IP addresses.
- **Trial number** – Even though trials are usually recorded in order, it is best to make this explicit.
- **One column for each independent variable** – It should be possible to reconstruct the trial.
- **A column for each raw measurement** – This means raw responses (e.g. whether the subject pressed left or right), not their accuracy or some other processed information. Aggregation or other processing is an important part of analysis that others might want to reuse or scrutinize.

Other suggested columns that may be useful depending on the specific experiment:

- **Subject information like gender, education, visual deficiencies, etc.** – It is most important for between-subject designs. Remember to be careful about anonymity.
- **Environment or equipment information like monitor resolution, browser, operating system, etc.**
- **Date and time** – When was the experiment run? When did each trial start?
- **Processed or aggregated information from other columns** – It is often useful to include processed information in the data. Just make sure that it augments rather than replaces raw information.

4.2 Keep it simple and tidy

If your raw data format is complicated, provide a reformatted copy of the data that is easy to process and analyze. The simplified format should still include all (or at least most) of the data, but it should be preprocessed for easier analysis. In other words, please simplify your arrays nested inside of JSON objects nested inside of CSV cells.

A format that's generally useful to analysis is the *tidy* approach [15] of putting all data in one table with each trial as a row and each column as a measurement. Also, you may want to modify condition names to be more readable if stored numerically (e.g., rename conditions 0 and 1 to the name of the value such as "red" and "blue"). Any additional improvements to facilitate reanalysis can be useful, even aggregation. However, it is critical that this processed data is a copy, not a replacement. Also provide the raw original data for transparency.

4.3 Use an accessible file type

Use free and open file types. Use to CSV when possible. JSON is acceptable if necessary, but it can be time-consuming to parse. If a project really needs some other format, make sure there are clear instructions for reading it. Providing code to read the data in common data processing languages (e.g., R and Python) can ameliorate the difficulties of complex file types.

4.4 Common mistakes

Aggregating the data – Some people post a single data point per subject or per condition despite having dozens of trials for each. Many assumptions are made when aggregating, so it is critical to provide raw unbiased data without locking people into a particular approach for aggregation.

- **Skipping the response variable** – While it is useful to know whether a response is correct, recording the actual response is more important in case there are concerns about how "correctness" was calculated.
- **Skipping the data dictionary** – Column names are often difficult for others to interpret. Make a text file with a very brief description of every column in your data.
- **Not putting the data URL in the paper** – How is anyone supposed to know how to get the data unless you put it in paper? Do not make anyone email you! I recommend putting it in the abstract.
- **Not using a reliable repository** – As section 1.2 already discussed, long term reliability and accessibility is key. Personal websites and repositories such as GitHub are not reliable and are easily modifiable. Tips for posting to OSF were mentioned in section 3, and other alternatives include Dryad Digital Repository, figshare, Harvard Dataverse Network, and Zenodo.
- **Failing to check text entries for identifying information** – You never know what information people will type into a textbox. One strategy is to drop that column from the open data and make it available on request.

4.5 Data size is rarely a concern

Open data repositories can handle large datasets. People manage to share huge results from astrophysical data to fMRI volumes that vary over time for dozens of subjects. A CSV or JSON that is under 5GB would easily fit on an open science repository like OSF and figshare. For larger datasets, you can break it up into multiple files or use repositories like Data Dryad.

5 When you cannot share

When sharing is not feasible, transparency is critical.

For qualitative experiments, raw data may include video or audio of subjects, and releasing it would violate their privacy. Similar privacy concerns could arise if patient health data is used in any capacity. In such cases, authors can explain in the paper that sharing the raw data would violate privacy, and the coded or aggregated data can be shared instead. As an alternative, a "Protected Access" notation exists for the open data badge wherein authors post data to a repository with restricted access that can be made available to researchers through a documented process (`https://osf.io/g6u5k/`).

Credit and ownership concerns can sometimes inhibit sharing especially when content comes from collaborators in industry or other areas of academia. For datasets and code generated by academics, credit and citations tend to be the primary reward of the collection effort. In such cases, clear citation policies for datasets, source code, and software may reduce the inhibition to share. The TOP guidelines offer example policies (see the first row of the appendix). In some cases, research groups may want to delay releasing a dataset until a high-impact long-term-project can be published. In such cases, data or source code can be posted to an embargoed repository, and the paper can state that the repository will be made public after some number of months or years.

Industry collaborators can be highly supportive of scientific research and help provide resources and funding without imposing restrictions on researchers' openness. However, intellectual property restrictions can sometimes limit the ability to share datasets, materials, and even results. Again, transparency is most critical, and the reasons for restricting access should be stated in the paper along with a clear explanation for how the work will still provide knowledge and understanding to the field. Reviewers can then use the explanation to make an informed decision. Importantly, reviewers should have the right to refuse a review or reject a submission if they are concerned that hiding resources prevents verification or that their review of a poorly generalizable application is merely a free consultation to a for-profit company. Peer-reviewed research publications are arguably the only literature where the validity of research is the primary priority, so openness and therefor the ability to scrutinize research should always take precedence over restrictive industry partnerships.

6 Further steps towards open practices in visualization research

A publication is not research; it is an editorialized summary of research results. The veracity of that summary can and should be questioned if the content being summarized is not available for scrutiny. For visualization research, the low rate of accessibility of supporting materials and data makes it almost impossible to question the claims made in publications. Everyone involved should actively support open practices as a way of demonstrating that the field's work is trustworthy.

Besides the necessary support of authors and the research community, the paper chairs and editors can play a critical role by adopting the Transparency and Openness Promotion guidelines, which provide a gradual mechanism of advancing openness with pre-written policy templates to avoid needing to "reinvent the wheel". Furthermore, adopting the open science badges would be a way of advertising this practice, informing readers from other fields that visualization researchers take transparency and accessibility seriously. Finally, in order to be consistent with the values of accessibility and application espoused by the visualization research community, the community should move towards an open access publication model. OAVIS has documented the state of open practices in visualization research, but it is up to the community as a whole to progress.

Acknowledgements

OAVIS would not be possible without the authors who shared their material and everyone who contributed to the crowdsourced data collection. Thanks to David Mellor and the Center for Open Science for helpful advice. Helpful feedback from Dominik Moritz was also appreciated.

This work was partially performed within the Labex SMART (ANR-11-LABX-65) supported by French state funds managed by the ANR within the Investissements d'Avenir programme under reference ANR-11-IDEX-0004-02.

Technical Acknowledgements

The OAVIS site was built on D3 (`http://d3js.org` [16]) and Bootstrap (`https://getbootstrap.com` [17]). Analyses and figure generation used R (`https://www.r-project.org/` [18]) and the Tidyverse (`http://tidyverse.org`).

References

[1] T. H. Vines, A. Y. K. Albert, R. L. Andrew, F. Débarre, D. G. Bock, M. T. Franklin, K. J. Gilbert, J.-S. Moore, S. Renaut, and D. J. Rennison, "The availability of research data declines rapidly with article age," *Curr. Biol.*, vol. 24, no. 1, pp. 94–97, Jan. 2014.

[2] B. A. Nosek and Y. Bar-Anan, "Scientific Utopia: I. Opening scientific communication," *Arxiv - http://arxiv.org/abs/1205.1055*.

[3] S. Serghiou and J. P. A. Ioannidis, "Altmetric Scores, Citations, and Publication of Studies Posted as Preprints," *JAMA*, vol. 319, no. 4, p. 402, Jan. 2018.

[4] H. A. Piwowar, R. S. Day, and D. B. Fridsma, "Sharing Detailed Research Data Is Associated with Increased Citation Rate," *PLoS ONE*, vol. 2, no. 3, p. e308, Mar. 2007.

[5] B. B. Blohowiak, J. Cohoon, L. de-Wit, E. Eich, F. J. Farach, F. Hasselman, A. O. Holcombe, M. Humphreys, M. Lewis, B. A. Nosek, J. Peirce, J. R. Spies, C. Seto, S. Bowman, D. Green, G. Nilsonne, J. Grahe, S. Wykstra, A. H. Mohr, A. Sallans, R. Giner-Sorolla, T. H. Parker, W. Forstmeier, S. Nakagawa, M. C. Kidwell, D. Mellor, and A. DeHaven, "Badges to Acknowledge Open Practices," Jun. 2018.

[6] J. M. Wicherts, C. L. S. Veldkamp, H. E. M. Augusteijn, M. Bakker, R. C. M. van Aert, and M. A. L. M. van Assen, "Degrees of Freedom in Planning, Running, Analyzing, and Reporting Psychological Studies: A Checklist to Avoid p-Hacking," *Frontiers in Psychology*, vol. 7, Nov. 2016.

[7] M. Kidwell, L. Lazarevic, E. Baranski, T. Hardwicke, S. Piechowski, L.-S. Falkenberg, C. Kennett, A. Slowik, C. Sonnleitner, C. Hess-Holden, T. Errington, S. Fiedler, and B. Nosek, "Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency," 2016.

[8] M. Kay, S. Guha, S. Haroz, and P. Dragicevic, "Special interest group on Transparent statistics in HCI," in *Conference on Human Factors in Computing Systems - Proceedings*, 2016.

[9] S. Haroz, "Experiment methods template." [Online]. Available: http://steveharoz.com/public/experimentmethods/.

[10] M. McNutt, "Taking up TOP," *Science*, vol. 352, no. 6290, pp. 1147–1147, Jun. 2016.

[11] T. Hardwicke, M. Mathur, K. MacDonald, G. Nilsonne, G. Banks, M. Kidwell, A. Hofelich Mohr, E. Clayton, E. Yoon, M. Tessler, R. Lenne, S. Altman, B. Long, and M. Frank, "Data availability, reusability, and analytic reproducibility: Evaluating the impact of a mandatory open data policy at the journal Cognition," 2018.

[12] R. D. Morey, C. D. Chambers, P. J. Etchells, C. R. Harris, R. Hoekstra, D. Lakens, S. Lewandowsky, C. C. Morey, D. P. Newman, F. D. Schönbrodt, W. Vanpaemel, E.-J. Wagenmakers, and R. A. Zwaan, "The Peer Reviewers' Openness Initiative: incentivizing open research practices through peer review," *Royal Society Open Science*, vol. 3, no. 1.

[13] B. L. Houtkoop, C. Chambers, M. Macleod, D. V. M. Bishop, T. E. Nichols, and E.-J. Wagenmakers, "Data Sharing in Psychology: A Survey on Barriers and Preconditions," *Advances in Methods and Practices in Psychological Science*, vol. 1, no. 1, pp. 70–85, Mar. 2018.

[14] C. K. Soderberg, "Using OSF to Share Data: A Step-by-Step Guide," *Advances in Methods and Practices in Psychological Science*, vol. 1, no. 1, pp. 115–120, Mar. 2018.

[15] H. Wickham, "Tidy Data," *Journal of Statistical Software*, vol. 59, no. 10, 2014.

[16] "D3.js - Data-Driven Documents." [Online]. Available: https://d3js.org. [Accessed: 18-Jul-2018].

[17] "Bootstrap · The most popular HTML, CSS, and JS library in the world." [Online]. Available: https://getbootstrap.com. [Accessed: 18-Jul-2018].

[18] R Core Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2018.

Appendix A: Transparency and Openness Promotion (TOP) summary. More information available at cos.io/top

Templates that can be used as journal or conference policies are available at https://osf.io/9f6gx/wiki/Guidelines/

	Not Implemented	Level I	Level II	Level III
Citation Standards	Journal encourages citation of data, code, and materials, or says nothing.	Journal describes citation of data in guidelines to authors with clear rules and examples.	Article provides appropriate citation for data and materials used consistent with journal's author guidelines.	Article is not published until providing appropriate citation for data and materials following journal's author guidelines.
Data Transparency	Journal encourages data sharing, or says nothing.	Article states whether data are available, and, if so, where to access them.	Data must be posted to a trusted repository. Exceptions must be identified at article submission.	Data must be posted to a trusted repository, and reported analyses will be reproduced independently prior to publication.
Analytic Methods (Code) Transparency	Journal encourages code sharing, or says nothing.	Article states whether code is available, and, if so, where to access it.	Code must be posted to a trusted repository. Exceptions must be identified at article submission.	Code must be posted to a trusted repository, and reported analyses will be reproduced independently prior to publication.
Research Materials Transparency	Journal encourages materials sharing, or says nothing.	Article states whether materials are available, and, if so, where to access them.	Materials must be posted to a trusted repository. Exceptions must be identified at article submission.	Materials must be posted to a trusted repository, and reported analyses will be reproduced independently prior to publication.
Design and Analysis Transparency	Journal encourages design and analysis transparency, or says nothing.	Journal articulates design transparency standards.	Journal requires adherence to design transparency standards for review and publication.	Journal requires and enforces adherence to design transparency standards for review and publication.
Study Preregistration	Journal says nothing.	Article states whether preregistration of study exists, and, if so, where to access it.	Article states whether preregistration of study exists, and, if so, allows journal access during peer review for verification.	Journal requires preregistration of studies and provides link and badge in article to meeting requirements.
Analysis Plan Preregistration	Journal says nothing.	Article states whether preregistration of study exists, and, if so, where to access it.	Article states whether preregistration with analysis plan exists, and, if so, allows journal access during peer review for verification.	Journal requires preregistration of studies with analysis plans and provides link and badge in article to meeting requirements.
Replication	Journal discourages submission of replication studies, or says nothing.	Journal encourages submission of replication studies.	Journal encourages submission of replication studies and conducts results blind review.	Journal uses Registered Reports as a submission option for replication studies with peer review prior to observing the study outcomes.

Source: https://osf.io/2cz65/