[The authors are currently seeking feedback on this preprint. Please email david@cos.io or comment directly on a working draft here: http://bit.ly/top2preprint ]

# Making Science Transparent By Default; Introducing the TOP Statement

The TOP Statement Working Group[1]

In order to increase the replicability of scientific work, the scientific community has called for practices designed to increase the transparency of research (McNutt, 2014; Nosek et al., 2015). The validity of a scientific claim depends not on the reputation of those making the claim, the venue in which the claim is made, or the novelty of the result, but rather on the empirical evidence provided by the underlying data and methods. Proper evaluation of the merits of scientific findings requires availability of the methods, materials, and data and the reasoned argument that serve as the basis for the published conclusions (Claerbout and Karrenbach 1992; Donoho et al 2009; Stodden et al 2013; Borwein et al 2013; Munafò et al, 2017). Wide and growing support for these principles (see, for example, signatories to Declaration on Research Assessment, DORA, https://sfdora.org/, and the Transparency and Openness Promotion Guidelines https://cos.io/our-services/top-guidelines/) must be coupled with

[1] Authors listed alphabetically: IJsbrand Jan Aalbersberg, Elsevier; Tom Appleyard, BioMed Central, Springer Nature; Sarah Brookhart, Association for Psychological Science; Todd Carpenter, National Information Standards Organization; Michael Clarke, Clarke & Company, STM Advisers; Stephen Curry, Imperial College; Josh Dahl, Clarivate Analytics; Alex DeHaven, Center for Open Science; Eric Eich, University of British Columbia; Maryrose Franko, Health Research Alliance; Len Freedman, Global Biological Standards Institute; Chris Graf, Wiley and COPE; Sean Grant, RAND Corporation; Brooks Hanson, American Geophysical Union; Heather Joseph, Scholarly Publishing and Academic Resources Coalition; Veronique Kiermer, Public Library of Science; Bianca Kramer, Utrecht University; Alan Kraut, Center for Open Science Board Member; Roshan Kumar Karn, Open Access Nepal; Carole Lee, University of Washington, Aki MacFarlane, Wellcome Trust; Maryann Martone, UC San Diego; Evan Mayo-Wilson, Johns Hopkins Bloomberg School of Public Health; Marcia McNutt, National Academy of Sciences; Meredith McPhail, Laura and John Arnold Foundation; David Mellor, Center for Open Science; David Moher, University of Ottawa, Equator Network; Alison Mudditt, Public Library of Science; Brian Nosek, Center for Open Science; Belinda Orland, American Heart Association; Tim Parker, Whitman College; Mark Parsons, Rensselaer Polytechnic Institute; Mark Patterson, eLife; Solange Santos, Scientific Electronic Library Online; Carolyn Shore, National Academy of Sciences; Dan Simons, Association for Psychological Science; Bobbie Spellman, University of Virginia; Jeff Spies, Center for Open Science; Matt Spitzer, Center for Open Science; Victoria Stodden, University of Illinois; Sowmya Swaminathan, Springer Nature; Deborah Sweet, Cell Press; Anne Tsui, Notre Dame, Simine Vazire, UC Davis

guidelines to increase open sharing of data and research materials, use of reporting guidelines, preregistration, and replication. *We propose that, going forward, authors of all scientific articles disclose the availability and location of all research items, including data, materials, and code, related to their published articles in what we will refer to as a TOP Statement.*

**The TOP Statement**

We propose that every scientific publication incorporates a statement that discloses the availability of those research products and materials that underlie its conclusions. The particular requirements are described in the Transparency and Openness Promotion (TOP) Guidelines: 1) the availability of data underlying empirical claims, 2) analysis methods or code used for data analysis, 3) physical or digital materials used to conduct the study, 4) research methods, described using community or disciplinary standards, or journal-created reporting guidleines (such as those created by Cell Press http://www.cell.com/star-authors-guide or Nature Research https://www.nature.com/authors/policies/ReportingSummary.pdf), and if applicable, 5) documentation of study registration (also referred to as "preregistration" in pre-clinical science), and 6) the scientific report. Obviously only a subset of items are relevant to each publication (e.g. a scientific report will of course not link to itself, though a published dataset would point to any relevant scientific report), and publishers can decide to include only the items that they deem most important to the work that they disseminate.

The disclosure would address a simple set of questions:

- Is the item publicly available in a persistent location? ("Yes", "Not applicable", or "See explanation")
- If Yes, provide unique, persistent identifier(s) and applicable license information.
- Otherwise, provide a brief explanation.

Individual components of the TOP Statement can be used by communities based on the standards desired within their discipline. Such modularity allows customization and reduces barriers to implementation. Some core items, such as data, materials, analysis code, and methods reports, are likely to be used most widely. This approach is highly achievable, as demonstrated by journals that currently implement similar data availability statements; the TOP Statement simply provides structure and recommended content to such statements. Suggested implementation strategies are provided below.

A template questionnaire for generating a TOP statement can be found at https://osf.io/n9mrh/, and an example of that template is provided for Kidwell et. al, (2016) here:

https://osf.io/zt4de/. Information included from TOP statements should be provided in a clear and consistent format, ideally in accessible, machine-readable, portable format so it is interoperable and can be ported to a number of different platforms. However, this ideal does not exclude other implementations, such as including the information within the article itself or in a separate file. Here is an example of a generic TOP Statement that uses language from existing data availability statements: https://osf.io/4mt26/.

Journals could also display the information included in TOP Statements for each article for readers, perhaps providing visual indicators of more open practices. If a service collected and displayed TOP Statements, it would benefit the scientific community by increasing the findability of items associated with any particular study, while making clearer how they can be accessed, and perhaps increasing the degree to which they are interoperable and reusable. These FAIR elements are important for increasing the overall transparency of scientific research (Wilkinson et al., 2016), and are compatible with existing policies already used by many journals, such as data availability statements or research materials disclosures.

A benefit of these minimal expectations for disclosure is that they are not subject to commonly cited barriers to more transparency into research, such as concerns about research participant confidentiality, location information for rare species susceptible to poaching, maintenance of competitive advantage, or intellectual property concerns. However, one important benefit of this minimal disclosure model is as a signal to policymakers, who will be able to quickly and accurately assess community norms in regards to persistent sharing of research products. It will allow the policy conversation to be based on facts gathered from near real time practices among researchers, as opposed to the opinions of any one stakeholder.

As envisioned, TOP Statements are a mandate for disclosure, not a mandate for access. We strongly advocate for additional solutions that would lead to more universal data sharing, but also recognize the difficulty or the cost of implementing them. The TOP Statement is a down payment toward more transparent and more reproducible research practices. By standardizing the practice of data and research materials disclosure, it will be easier to implement further actions.

**Implementation of the TOP Statement**

TOP Statements can be associated with any research output, not just a journal article, and so can be supported by organizations that publish, store, or fund research outputs. Furthermore, even adoption by a subset of stakeholders will provide benefits to the entire community.

Standalone journals could implement a TOP Statement into the article publication process by requiring the authors to include the information covered by the TOP statement in their paper, or by asking authors to complete a checklist prior to publication, or by requiring a TOP Statement as part of the article. Publishers could implement TOP Statements directly into the article submission process of many journals at once, which would facilitate more rigorous review standards for those who are able to implement them. Wide adoption of TOP Statements would simplify interoperability, increase discoverability, and facilitate portability from one journal to another. Similarly, data and code repositories could request TOP statements from those using their platforms, which would facilitate the discovery of articles derived from the data. Publishers that have implemented checklists in order to increase transparency and comprehensiveness have seen some improvements in reporting (Macleod, 2017).

Funders could request TOP Statements as part of their existing workflow with researchers, for example, during interim or final progress reports. However, some of the benefits of TOP Statements to funders will be achievable even with low uptake by the funding community, as TOP Statements could still be used to evaluate a researcher's record of transparency even if they are only attached to publications and not to funding agreements.

**Benefits of TOP Statements**

Everyone stands to benefit from increased transparency about the process of science, and the purpose of the TOP Statement is to address the existing collective action problem. This collective action problem exists because although the scientific community benefits by increased transparency into individual studies, current norms (and incentives) mean that individual researchers typically only provide as much transparency as is necessary for publication. Increased transparency is difficult to both evaluate and reward. TOP statements will provide an opportunity to improve on current norms by facilitating a more complete and long-lasting understanding of the connected network of materials underlying the research, and thereby extending the reach of a paper beyond a fixed moment in the lifecycle of a research project.

Standard use of TOP statements would have wide ranging benefits. Researchers will be able to more easily adopt, test or build upon the work of others. Authors will benefit through increased credibility or understanding of their results, the potential for more re-use, citation, and recognition of their outputs. Peer reviewers will benefit by being more able to critically evaluate a claim. Journals will benefit from uniformity in how information about transparency is conveyed through, and a better connection to, the research that underlies their publications. Universities

will have more accurate knowledge about the work conducted by their researchers and will be able to take the transparency and reproducibility of a scientist's work into account for hiring and promotion decisions (e.g. as recommended by Flier, 2017). Funders will know the extent to which the research and resources they supported are being made available to fellow scientists, policy makers, or consumers of scientific knowledge. Patients will benefit from increased transparency into the evidence base that informs their treatment. Society will benefit from more reliable research, increased access to the basis for scientific conclusions, and more accessible information to guide policy. The research community at large will benefit from increased public trust in its work. Each stakeholder wants and needs this basic information. Requesting it of everyone will address the collective action problem inherent in the culture change required to increase overall transparency.

These goals can be achieved through common understanding of what items must be disclosed that support any empirical scientific claim and through clarifying minimal expectations for the research community.

**Enabling a future of more transparent research**

By adopting a TOP Statement approach for research outputs, an organization such as a publisher not only asserts its principled support for increasing transparency into the process of science, but also articulates and implements its policy expectations, *i.e.* that disclosing availability of key elements of research, such as data, is required. Though there are justifiable barriers to universally sharing all data, such as legal constraints or promises of confidentiality made to research participants, researchers can still disclose and justify what is available.

Implementing more rigorous policy standards, such as mandates for transparency or verification of shared outputs, also becomes more achievable once TOP Statements become common practice. TOP Statements lay the necessary groundwork for future, more rigorous policies by standardizing where relevant information or metadata would be located. Then, implementing these policies becomes a matter of evaluating any given TOP Statement according to the new requirement. Evaluating the justification for unshared data or materials becomes the task of those implementing and enforcing any new policy.

Evaluating adherence to a policy also becomes less burdensome after implementing a TOP Statement. TOP Statements become a resource to those who are comparing stated policy to actual practice. Indeed, the mere use of the TOP Statements guarantees compliance with disclosure requirements. Other policies, such as mandated public deposition of data, would be

easier to enforce if the necessary information were provided in a standardized way that is facilitated by the use of TOP Statements.

Where transparency is not possible, as is the case with very large datasets or where ethical constraints prohibit it, disclosure is nevertheless still valuable in order to provide more context of that data. Furthermore, TOP Statements need not apply only to the typical article as we know it today: connections to related content can and should be made from preprints, published datasets, grant applications, and other research outputs. However, their adoption will be a commonsense step that will make science more transparent. Meanwhile, TOP Statements will allow each stakeholder the option to continue to determine the policies that are ideal for their community.

The wide adoption of TOP Statements will not be a panacea for all that ails science. Implementation will require effort and creative problem solving by researchers, journals, publishers, and funders to establish new norms for research publication. The investment in time and money will have been worthwhile if it increases the credibility, trust, and reproducibility of scientific findings. Increasing transparency into the process of science is critically important to the trust that society holds in its institutions, and we commit to making the above the minimal default for any research output.

---

**Literature cited**

Alsheikh-Ali, A. A., Qureshi, W., Al-Mallah, M. H., & Ioannidis, J. P. A. (2011). Public Availability of Published Research Data in High-Impact Journals. PLoS ONE, 6(9), e24357. https://doi.org/10.1371/journal.pone.0024357

Caron, J. E., March, J. K., Cohen, M. B., & Schmidt, R. L. (2017). A Survey of the Prevalence and Impact of Reporting Guideline Endorsement in Pathology Journals. American Journal of Clinical Pathology, 148(4), 314–322. https://doi.org/10.1093/ajcp/aqx080

Claerbout, J., & Karrenbach, M. (1992). Electronic documents give reproducible research a new meaning. In SEG Technical Program Expanded Abstracts 1992 (Vols. 1–0, pp. 601–604). Society of Exploration Geophysicists. https://doi.org/10.1190/1.1822162

Donoho, D. L., Maleki, A., Rahman, I. U., Shahram, M., & Stodden, V. (2009). Reproducible Research in Computational Harmonic Analysis. Computing in Science & Engineering, 11(1), 8–18. https://doi.org/10.1109/MCSE.2009.15

Flier, J. (2017). Faculty promotion must assess reproducibility. Nature, 549(7671), 133–133. https://doi.org/10.1038/549133a

Hicks D, Wouters P, Waltman L, de Rijcke S, Rafols I. Bibliometrics: The Leiden Manifesto for research metrics. Nature. 2015;520(7548):429–31

Kvalheim, V., & Kvamme, T. (2014). Policies for Sharing Research Data in Social Sciences and Humanities. International Federation of Data Organizations. Retrieved from http://ifdo.org/wordpress/ifdo-report-2014-policies-for-sharing-research-data-in-social-sciences-and-humanities/

Macleod, M. R. (2017). Findings of a retrospective, controlled cohort study of the impact of a change in Nature journals' editorial policy for life sciences research on the completeness of reporting study design and execution. https://doi.org/10.1101/187245

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in Psychology Research: How Often Do They Really Occur? Perspectives on Psychological Science: A Journal of the Association for Psychological Science, 7(6), 537–542. https://doi.org/10.1177/1745691612460688

Martin, G. N., & Clarke, R. M. (2017). Are Psychology Journals Anti-replication? A Snapshot of Editorial Practices. Frontiers in Psychology, 8. https://doi.org/10.3389/fpsyg.2017.00523

McNutt, M. (2014). Journals unite for reproducibility. Science, 346(6210), 679–679. https://doi.org/10.1126/science.aaa1724

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Sert, N. P. du, … Ioannidis, J. P. A. (2017). A manifesto for reproducible science. Nature Human Behaviour, 1, 0021. https://doi.org/10.1038/s41562-016-0021

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., … Yarkoni, T. (2015). Promoting an open research culture. Science, 348(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

Rigor and Reproducibility | grants.nih.gov. (n.d.). Retrieved January 5, 2018, from https://grants.nih.gov/reproducibility/index.htm

Stodden, V., Guo, P., & Ma, Z. (2013). Toward Reproducible Computational Research: An Empirical Analysis of Data and Code Policy Adoption by Journals. PLoS ONE, 8(6), e67111. https://doi.org/10.1371/journal.pone.0067111

Vogt, L., Reichlin, T. S., Nathues, C., & Würbel, H. (2016). Authorization of Animal Experiments Is Based on Confidence Rather than Evidence of Scientific Rigor. PLOS Biology, 14(12), e2000598. https://doi.org/10.1371/journal.pbio.2000598

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018. https://doi.org/10.1038/sdata.2016.18

---

Competing interests

Please see the author affiliations to understand the perspective that they bring to the table. Of course, being an author on this paper does not imply that the represented organization has endorsed any particular aspect of this document.

Five authors (Alex DeHaven, David Mellor, Brian Nosek, Jeff Spies, and Matt Spitzer) are employed by the Center for Open Science, which builds and maintains the free and open source OSF.

Veronique Kiermer is chair of ORCID Board.