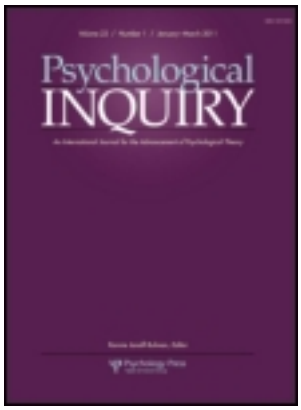
This article was downloaded by: [Karen E. Adolph]
On: 12 September 2012, At: 03:00
Publisher: Psychology Press
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Psychological Inquiry: An International Journal for the Advancement of Psychological Theory

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hpli20

Toward Open Behavioral Science

Karen E. Adolph [a], Rick O. Gilmore [b], Clinton Freeman [c], Penelope Sanderson [c] & David Millman [d]

[a] Department of Psychology, New York University, New York, New York

[b] Department of Psychology, The Pennsylvania State University, University Park, Pennsylvania

[c] Department of Psychology, The University of Queensland, St Lucia, Queensland, Australia

[d] Division of Libraries, New York University, New York, New York

Version of record first published: 10 Sep 2012.

To cite this article: Karen E. Adolph, Rick O. Gilmore, Clinton Freeman, Penelope Sanderson & David Millman (2012): Toward Open Behavioral Science, Psychological Inquiry: An International Journal for the Advancement of Psychological Theory, 23:3, 244-247

To link to this article: http://dx.doi.org/10.1080/1047840X.2012.705133

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Psychological Inquiry, 23: 244–247, 2012
Copyright © Taylor & Francis Group, LLC
ISSN: 1047-840X print / 1532-7965 online
DOI: 10.1080/1047840X.2012.705133

Psychology Press
Taylor & Francis Group

Downloaded by [Karen E. Adolph] at 03:00 12 September 2012

COMMENTARIES

Toward Open Behavioral Science

Karen E. Adolph
Department of Psychology, New York University, New York, New York

Rick O. Gilmore
Department of Psychology, The Pennsylvania State University, University Park, Pennsylvania

Clinton Freeman and Penelope Sanderson
Department of Psychology, The University of Queensland, St Lucia, Queensland, Australia

David Millman
Division of Libraries, New York University, New York, New York

Brian A. Nosek and Yoav Bar-Anan (this issue) describe three traditional ways of doing science: Researchers can operate independently, antagonistically, or collaboratively. There is also an emerging fourth way: Scientists can operate openly. In their vision of scientific utopia, Nosek and Bar-Anan call for increased openness in peer review and journal access. By opening up the review process and providing everyone with free digital access to published papers, the process of communicating and disseminating information would become a continuous conversation rather than a monologue delivered into the void of unpublished manuscripts or a dialogue separated by interminable periods of silence between submission, review, publication, and subsequent commentaries and follow-up studies. With open review and faster access, more people could be part of the process and credited with their contributions, and science would benefit.

Other areas of scholarship stand to benefit from openness. The MediaCommons Network family of digital-only journals in Media, Culture, and Communications Studies (mediacommons.futureofthebook.org) encourages a rich and public review process. Scholars in the humanities are beginning to invite public review of monographs prior to print publication (Cohen, 2010). Library scientists at New York University, the University of Illinois, and elsewhere are developing technical standards that permit complex scholarly commentary structures, such as group review and hierarchical review conversation.

We have a complementary vision of utopian openness—open data sharing, where researchers contribute data to a secure repository and can browse and download deposited files of interest. Open data sharing is especially important in developmental science where data from infants, young children, and special populations are difficult and expensive to collect, and many longitudinal and clinical studies contain unique data that cannot be readily replicated. A primary opportunity for data sharing among developmental scientists concerns the widespread reliance on video recording. Most developmental psychologists, clinicians, therapists, and movement scientists rely on video as their primary data collection method because it is low cost, is easy to use, and faithfully renders the richness and complexity of behavioral development. However, the wealth of analytic possibilities remains largely untapped because developmental scientists do not mine the richness of their own video data or engage in open sharing of video files. After publication, the largely underused videos molder quietly on a shelf. With open video data sharing, substantial public investments in behavioral research could yield their full potential.

Are these utopias of openness possible? With regard to open access and peer review, Nosek and Bar-Anan point to the notable success of arXiv (arxiv.org) and the Social Science Research Network (ssrn.com) that provide researchers with open access to manuscripts prior to publication in physics, math, cognitive science, economics, and other scientific communities. In addition, open sharing of raw data is common practice in the physical, biomedical, and earth sciences, and it is growing in the biological and social sciences. Examples of sharing repositories in neuroscience include the Biomedical Informatics Research Network (www.birncommunity.org) and the International Electrophysiology Epilepsy Data Portal (braintrust.seas.upenn.edu/); examples in the behavioral/social

sciences, include the Inter-university Consortium for Political and Social Research (www.icpsr.umich.edu/), TalkBank and CHILDES (talkbank.org), the Social Informatics Data Grid (sidgrid.ci.uchicago.edu/home/), and the Digital Interactive Video Exploration & Reflection (DIVER; diver.stanford.edu) project.

Thus, open data sharing in the behavioral sciences has seen preliminary success, whereas in other areas of science, researchers are already living the dream. However, existing data repositories support open sharing only of text and image files. Open sharing of video data is the next big challenge. Certainly, digital video poses technical problems for sharing and raises ethical issues for protecting participants' privacy, but overcoming these challenges will broaden the generalizability of findings, facilitate peer oversight, and speed progress.

Downloaded by [Karen E. Adolph] at 03:00 12 September 2012

Impact, Transparency, and Truth

Why share? One good reason to share data is that everyone—researchers, funders, the public—gets a bigger bang for their buck. Data sharing accelerates discovery. Much more rapid progress would be made if researchers could build on earlier efforts by analyzing videos in ways unimagined by the original researcher; if researchers could browse for examples to stimulate new work or to train and educate students; and if researchers could gather preliminary data, expand samples, run replications, examine cohort effects, and assess effects of geographic location or population using data in a shared archive. Researchers could use tools contributed to a shared repository to enhance understanding of their own data and use shared data to test their analytic tools. They could collaborate with like-minded researchers in subareas of interest to create corpora with shared coding schemes and analysis tools (e.g., TalkBank). A history of data sharing and a plan to commit data to an open repository would enhance the likelihood of Federal funding. Researchers' work would receive more attention and citations by users, and their data would survive in useable form beyond their lifetimes. These are not our claims, although we endorse them. They come from the NIH Data Sharing Workbook (National Institutes of Health, 2004) and the NSF 2011 National Science Board report on data sharing (National Science Foundation, 2011).

Additional motivations for open science are transparency and truth. Nosek and Bar-Anan point out that openness is a central scientific value. Given the distributed nature of scientific practice, open review, open access, and open data sharing increase efficiency and veracity in the construction and dissemination of scientific knowledge. Recent reports of data fraud have threatened the integrity of behavioral science; open data sharing would discourage practices that lead to the manufacture and fraudulent manipulation of data.

Although fraud is rare, a pervasive problem is incomplete communication. With our current system, Method sections in published articles cannot report everything that was done and Results sections cannot report everything that was found. The enforced brevity of our published communications makes it impossible for readers to fully evaluate the rationale and validity of researchers' coding decisions and analyses. We could lay a much firmer foundation if there were true transparency such that researchers could literally view each other's methods and phenomena.

Here again, video provides an avenue for openness (Gelman, 2012). Access to raw video data would reveal the entire experimental protocol, including the physical and social context, stimuli, spatial layout, experimenter language and intonation, gaze, pacing, gestures, feedback, and so on. It would also reveal the raw phenomena in their entirety and the links between data, metadata, and conclusions. Video captures more of the detail and nuance than a brief verbal description, photo, or diagram, and it provides clearer bases for replication and review. Video also provides for deeper, richer understanding of the phenomena. If a picture is worth a thousand words, a video is worth many thousands of pictures.

Community

Nosek and Bar-Anan argue that open review engages and strengthens the research community. By formally acknowledging reviewers and making their hard work public and part of the continuous research conversation, researchers would be more motivated to review each other's work and more likely to reap the benefits of feedback before their work has progressed to the point of no return. As Nosek and Bar-Anan suggest, those scientists who are particularly thorough, thoughtful, and expert in their commentaries about others' research could put those important talents to work and be acknowledged for their contributions as "review experts." For behavioral scientists, open video data sharing adds an important ingredient to open peer review—access to the raw data. Researchers would have the option of pointing reviewers to video files illustrating the methods and results. Reviewers would have the option of seeing the raw data—the methods and phenomenon—for themselves and of checking an author's analyses. Peer review could only benefit from access to the data.

Current practices in open source software development amplify these points. Continuous peer review is a core feature of the open source software community. Peers can see exactly what you did, how you did it, and how well it works. Many tools support discussions about source code, such as the review tool in github

(https://github.com/features/projects/codereview). Peers review code as it is being written, not only after the software is released, and thus offer guidance from the earliest stages of the project's development. This saves costly reworking. Openness allows failure to emerge earlier and in public. This provides rapid, actionable feedback and helps to avoid lengthy delays in development. Openness improves the quality and skill of individual developers because as they struggle to solve a particular kind of problem, they can consult the community to learn how others have solved similar problems. Open, continuous review enables software developers to learn from their own mistakes and the mistakes of others much faster than the traditional review process.

Openness leads to inclusiveness and cross-fertilization. It pulls people in and grows the community. Investigators can make substantive contributions through reviews, theoretical critiques, and tool development and benefit from a system that credits and acknowledges those contributions. Similarly, scientists from other disciplines can share their expertise by contributing reviews, theoretical frameworks, or analysis tools that can be applied to behavioral and developmental science. Open video data sharing is especially ripe for cross-fertilization within and across disciplines. Researchers with limited financial, technical, or practical resources—perhaps just lacking the know-how for data collection—could use previously collected data for their own purposes. Early stage investigators, for example, could compile pilot data from the repository. Computer scientists could analyze the video data for the purpose of studying scene and pattern recognition. Mathematicians and statisticians could use the data for modeling. Of importance, in all cases, the PIs who originally collected the data would be cited and acknowledged. Creating a culture of openness in the behavioral sciences would simultaneously improve the quality of research and make the process of discovery more inclusive, interdisciplinary, and efficient. This vision, however, requires a sea change in current practice: community buy-in to the vision of openness.

Downloaded by [Karen E. Adolph] at 03:00 12 September 2012

Barriers and Solutions

In Nosek and Bar-Anan's view, the primary barrier to openness is social. The scientific community is the primary impediment to change and therefore must be the instigator of change. Researchers' inertia and uncertainty impede progress toward open review and open access. The same holds true for open data sharing. According to the directors of existing efforts in the behavioral and neural sciences, open data sharing works only when it has grassroots support from the community of active researchers (G. Alters, personal communication, February 10, 2012; Bertenthal, 2011; C. Kesselman, personal communication, February 8, 2012; B. MacWhinney, personal communication, September 16, 2011). Advocates for greater openness must engage in the hard work of forging consensus and meeting varied researcher needs.

We recently mounted a comprehensive effort to galvanize the developmental science community toward openness in research (databrary.org). Our focus has been on identifying the barriers to open sharing of video files and associated metadata (openshapa.org/share/nsf-openshare-workshop-summary.pdf) and finding concrete and meaningful ways to address them. All of our efforts have been and will continue to be informed by peer input. In developing our ideas for the Databrary.org project, we reached out to colleagues and actively sought their input and ideas. Already, more than 100 developmental scientists (see databrary.org) have endorsed a vision of open research and have committed to seeking permission to share video data. So, a critical mass of the developmental science community appears poised to embrace open video data sharing—if the practical barriers to doing so can be effectively overcome.

In contrast to Nosek and Bar-Anan's claim that technical and financial concerns are not important barriers to open review and open access, open video data sharing does involve substantial technical and financial barriers. It poses ethical barriers, as well. Video cannot be anonymized in the same way as text files and neural images. Video coding—applying user-defined "tags" to events of interest—is inherently idiosyncratic and study specific. Video files and associated metadata must be encoded into common, sustainable formats. Video files are huge, such that storage, browsing, and streaming are technically challenging. All of this costs money, as does the development of suitable software and establishing the appropriate cyber infrastructure. Thus, the barriers to open sharing of video data include privacy/permission issues and technical/financial challenges, as well as entrenched community practices of isolation. The Databrary.org project seeks to overcome these challenges.

We are tackling the participant privacy/permission challenge by building on the work of previous data sharing efforts (e.g., TalkBank, ICSPR) so as to guard against inappropriate use of contributed data. We are creating standard templates for participant permission, data use, and data contribution, with extensive input from the research community and experts in research ethics with human participants. These efforts build on the common practice of seeking participant permission to show video excerpts in research contexts such as scientific meetings. We have posted our draft sharing permission template at databrary.org.

The technical challenges of sharing video files and associated metadata are myriad and diverse. To meet the problem of a common format for tagging and exploring video files, we have released a free, open source video coding tool, OpenSHAPA (openshapa.org) and the software continues to be developed with critical input from developmental and behavioral scientists who use it in their research. The success of OpenSHAPA to enable open video sharing requires interoperability with other academic and commercial video coding software. Developing tools for data interoperability is a central goal of the Databrary.org project. Another technical barrier is that none of the existing repositories for social science data (ICPSR, TalkBank) have the capacity to store large quantities of video data (100s of TB) or the software to enable data discovery or permit data mining from video. Thus, we have proposed to create a high-capacity repository and data exploration system that permits developmental scientists to share and reuse video and the associated metadata. The technical barriers to open video sharing are difficult but surmountable.

In conclusion, openness is good for science. Open review, open access, and open data sharing are imperative for scientific progress. Sharing video is central to greater efficiency and transparency in the developmental and behavioral sciences. The developmental science community is ready to embrace open video data sharing, and we have concrete and practical ways to address many of the barriers. Nevertheless, it is clear from other areas of science that have already adopted open practices that open data sharing requires a substantial, long-term financial commitment from federal funding agencies, private foundations, and research institutions. We applaud Nosek and Bar-Anan for their instructive analysis of the benefits of greater openness in scientific research and look forward to a continuous—and we hope transformative—conversation about openness with our peers.

Acknowledgments

This work was supported by National Institute of Health and Human Development Grant R37-HD33486 to Karen E. Adolph and National Science Foundation award 1147440 to Rick O. Gilmore.

Note

Address correspondence to Karen E. Adolph, Department of Psychology, New York University, 6 Washington Place, Room 410, New York, NY 10003. E-mail: karen.adolph@nyu.edu

References

Bertenthal, B. I. (2011). *Cyberinfrastructure for collaborative research in the neural, social, and behavioral sciences*. Retrieved from http://openshapa.org/news/archives/440

Cohen, P. (2010, August 24). Scholars test web alternative to peer review. *The New York Times*. Retrieved from http://www.nytimes.com/

Gelman, S. A. (2012). Technology could help. *APS Observer*. Retrieved from http://www.psychologicalscience.org/index.php/publications/observer/scientific-rigor.html#gelman

National Institutes of Health. (2004, February 13). *Data sharing workbook*. Bethesda, MD: Author. Retrieved from http://grants.nih.gov/grants/policy/data_sharing/data_sharing_workbook.pdf

National Science Foundation. (2011, December 14). *Digital research data sharing and management* (NSB-11-79). Arlington, VA: Author. Retrieved from http://www.nsf.gov/nsb/publications/2011/nsb1124.pdf

Downloaded by [Karen E. Adolph] at 03:00 12 September 2012