Data Sharing: Principles and Considerations for Policy Development

Data Sharing Working Group*
American Psychological Association

June 2015

The American Psychological Association's Board of Scientific Affairs (BSA) convened a working group in 2014-15 to explore the ethical and practical issues involved in sharing research data. In addition to members of BSA, the working group included representatives from other APA governance bodies: the Committee on Animal Research and Ethics, the Committee on Human Research, and the Publications and Communications Board.

PRINCIPLES

The working group affirmed the following principles:

- Sharing data promotes scientific progress.
- Sharing data within the larger scientific community encourages a culture of openness and accountability in scientific research.
- Sharing data allows geographically dispersed individuals and those with limited resources to investigate scientific questions of interest, enables replication of analyses for verifying empirical findings, and opens extant data to analysis with new, more powerful, or integrative techniques than were available at the time of collection.
- Sharing data promotes aggregation for the purposes of knowledge synthesis, hypothesis generation, programmatic decision-making, and generalizability testing, which ensures that the value of data is maximized.

This document is intended to be a contribution to broader discussions on data sharing policies. It does not constitute APA policy or commit APA to any particular activities.

CONSIDERATIONS

Policies regarding data sharing are currently being developed by government agencies, research and academic institutions, publishers, scientific societies, and other entities. The working group offered the following considerations to guide development of data sharing policies. Although the focus is on data within the behavioral and social sciences, most of these considerations apply across all scientific disciplines.

1. Formulation of data sharing policies requires careful balancing of rights and responsibilities, and of benefits and burdens, across research participants, the public, researchers, and institutions.

2. Data sharing policies should define specific levels or conditions for access, ranging from fully open access to limited access with permission.

3. Appropriate access levels, methods, formats, and timing of data sharing vary with the type of data collected. For example, procedures for sharing data that have the potential for identifying individual human participants, violating confidentiality, or identifying sites of illegal or stigmatized behavior need to be carefully designed and monitored. Of particular concern is the sharing of qualitative data (e.g., transcripts of interviews, including clinical interviews), geographical data (such as from GIS), field and research notes, biological data (including specimens), and photographic, video, and other image data.

4. Timing requirements for data sharing (i.e., how long researchers can wait to share data that they produced) should be commensurate with the research team's investment of effort in study conceptualization and implementation, as well as with the time required for the research team to conduct its own analyses of the data.

5. Sharing of data must conform to applicable statutes and regulations and to prior agreements with other parties, including participants' consent and—in the case of community-based participatory research—agreements concerning community ownership or control of the data. For international research, policies and procedures for sharing data should be sensitive to the regulations, norms, and other social-cultural characteristics of the settings in which the data are collected.

6. It is the responsibility of the principal investigator to ensure that data are prepared in a form that is usable and interpretable by others (including provision of sufficient background information and annotations) and to deposit the data and accompanying material in a repository in accord with applicable data sharing policies. The secondary data user is responsible for protecting the identities and confidentiality of participants, conducting analyses consistent with best methodological practices, making informed and responsible

interpretations of the data, and consulting with the principal investigator or original research team as needed to ensure that the user has proper understanding of the data.

7. Research and academic institutions and scientific publishers should establish standards for data management and sharing and for storage and preservation of data in secure repositories. Such standards should promote the valid use and interpretation of data. Current publication practices that may hinder users' full understanding of how data were collected, processed, or analyzed—such as word limits and restrictions on provision of online supplementary material—should be re-examined. In addition, standards need to be developed for citing and crediting the original producers of data in publications authored by secondary users of the data.

8. Research and academic institutions and scientific societies should sponsor training and educational programs for researchers and students that address technical, policy, and ethical aspects of data sharing.

9. The criteria that research and academic institutions use to evaluate researchers and to make promotion decisions should incorporate acknowledgement of the value of data sharing and of secondary uses of shared data.

10. An appropriate distribution of the work and costs underlying data sharing needs to be developed among researchers, funding agencies, research and academic institutions, publishers, and other stakeholders.