

## Guidelines for Researchers on Ethical Issues around Open Access Data

There is increasing impetus (from funders and the scientific community more generally) for the data collected as part of research projects to be made available to other researchers and perhaps more broadly, via for example, the Open Science Framework (<https://osf.io>) or the University's online research data archive (<https://orda.shef.ac.uk>). Furthermore, the University's policy note on the reuse of existing data states that "*all researchers are strongly encouraged to consider the possibility of secondary research and data sharing at the outset, before the primary data collection begins, and to build this in to the informed consent process.*" (<https://www.sheffield.ac.uk/ris/other/gov-ethics/ethicspolicy/policy-notes>). With this in mind, members of the Department of Psychology Ethics Committee (DESC) met in August 2017 in an effort to provide guidance to researchers in considering these issues.

In general, the Department of Psychology supports open science and the associated sharing of research data with the scientific community and more broadly. However, we recognise that this must be done in a manner that protects those who participated in that research. Therefore, our overarching principle is that **data can be shared so long as it is completely anonymous such that individual respondents cannot be identified**. Note that this principle can be taken to apply retrospectively, even if participants did not explicitly consent to their (anonymised) data being shared before taking part in the research.

This overarching principle does, however, come with a couple of caveats. First, please be aware that anonymising the dataset (e.g., removing names, email address, postcodes, or other identifying information) does not necessarily mean that participants cannot be identified. For example, if your research focuses on participants with a very rare genetic condition or elite level athletes, then it may be possible to guess who took part in the research. Second, please consider the extent to which procedures designed to anonymise the data (e.g., distorting audio files, or pixelating videos) actually means that participants cannot be identified. In general, DESC is not keen on the open access sharing of audio or visual data, even if participants have provided explicit consent for this, as they may not appreciate the implications of having audio/visual information in the public domain.

Furthermore, we recommend that going forward, participants (or the person providing consent for the research to take place - e.g., a headteacher at a school) are also made aware and consent to their data being used by other researchers and deposited in open access data repositories - in an anonymous fashion, such that their individual responses cannot be traced. In doing this, it may also be appropriate to explain the rationale for sharing data (e.g., to aid scientific collaboration and inquiry).

Finally, please note that Dr Tom Stafford represents the Faculty of Science on the 'scholarly communication' advisory group and so would be happy to discuss issues around open access data with researchers, if needed <https://www.sheffield.ac.uk/psychology/staff/academic/tom-stafford>

Thomas Webb, Emma Blakey, Megan Freeth, Anita Kenny, Liat Levita, and Tom Stafford, August 2017