

# METHODOLOGY CENTER PROFESSIONAL CONDUCT GUIDELINES

## Guidelines document 3: DATA SHARING

### 3.1 The importance of data sharing to The Methodology Center

Empirical data play a vital role in the research of this Center. Most of our publications involve analyses of empirical data. Thus, maintaining access to high-quality data sets that are relevant to our work is tremendously important.

### 3.2 Where do data come from?

Most of the scientists in our Center never collect data in the field. Instead, we depend on colleagues who graciously share their data. As a result, for most of us “data collection” means a 20-second download from a web site or receipt of a CD in the mail.

As a result, we may underestimate the time, effort, money, teamwork, expertise, dedication and stamina required to plan and execute collection of good social and behavioral science data, particularly longitudinal data. Anyone who makes this investment to collect data does so to address a set of important scientific questions and expects to base numerous publications on the data set. Such individuals should not be expected simply to hand over the data to someone else (i.e. one of us) with no strings attached.

For this reason the Center strives to maintain clear agreements spelling out the rights of those who provide the data, the responsibilities of Center scientists who analyze the data, and expectations on both sides about how collaboration will proceed.

### 3.3 Data sharing agreements

With the exception of public domain data, every data set available for use by Center scientists must be associated with a data sharing agreement. Usually this agreement will be a tailor-made version of our in-house data sharing agreement (see the Sample Data Sharing Agreement on the Center’s web site). This agreement spells out the rights and responsibilities of both the Center and the owners of the data. In some cases the owners of the data have their own data sharing agreements, which we observe.

A data sharing agreement is on file for every non-public data set in use in the Center (see the Scientific Director). It is your responsibility to become familiar with the data sharing agreements corresponding to data sets you analyze and to abide by their terms. Failure to do so can result in temporary or permanent loss of access to Center data.

### 3.4 Data sets not currently archived by the Center

If you wish to use a data set not in the public domain and not currently archived by the Center, such as a data set offered to you personally by a colleague outside the Center, a

data sharing agreement is necessary in every case (even if the owner of the data says that no agreement is needed). The agreement can be based on our in-house data sharing agreement or on terms stipulated by the owners of the data. See the Center's Scientific Director for help with this task. An original, signed version of the data sharing agreement must be stored in the Center archives.

### **3.5 Use of data in written documents produced by Center scientists**

If you use a data set in a written document of any kind, you must give the owners of the data an opportunity to be an author on the document, unless (a) the owners of the data have waived this right AND this is documented in writing, or (b) the data set is in the public domain. This policy holds even for incidental examples in primarily technical papers. An owner of data who declines the opportunity to be an author should be thanked by name in the acknowledgments.

Also see "Keeping co-authors in the loop" in the Authorship (Number 4) document in this series.

Usually a federal or private grant supported the original data collection. This grant should be mentioned in the acknowledgments by number or title as appropriate. (Federal grants should always be referenced by NIH number.) The first author must check with the owners of the data to determine the acknowledgements they require.