



Council on Undergraduate Research

President

Michael Nelson

Univ. of Wisconsin – La Crosse

Immediate Past President

Toufic Hakim

Kean University

President Elect

Mitchell Malachowski

University of San Diego

Treasurer

John P. Idoux

Tarleton State University

Secretary

Michael Tannenbaum

Marist College

At-Large Division Chair

Rosalind R. Reichard

Meredith College

Biology Division Chair

Jeffrey Osborn

Truman State University

Chemistry Division Chair

Kerry Karukstis

Harvey Mudd College

Geosciences Division Chair

Jill K. Singer

SUNY – College at Buffalo

Mathematical & Computer Sciences Division Chair

Patrick MacDonald

New College of Florida

Physics/Astronomy Division Chair

Jeffrey Collett

Lawrence University

Psychology Division Chair

MaryAnn Baenninger

Middle States Commission on Higher Education

Social Science Division Chair

Elizabeth Paul

The College of New Jersey

Ex-Officio:

Mark Shapiro

California State University - Fullerton

Dr. K. Elaine Hoagland

National Executive Officer

July 6, 2001

Brooke Dickson

Office of Information and Regulatory Affairs

Office of Management and Budget

Washington, DC 20503

Dear Ms Dickson:

The Council on Undergraduate Research, representing over 850 colleges and universities across America with a mission to promote integration of research and undergraduate education, would like to comment on Federal Register Notice: June 28, 2001, Vol. 66, No. 125, pp 34489-34493, "Proposed Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies". This notice relates to implementation of Section 515 of the Treasury and General Government Appropriations Act for FY 2001 (PL 106-554).

The Council On Undergraduate Research is concerned that this proposed rule, intended to ensure objectivity and integrity of information, would do just the opposite by allowing, even requiring, the tampering of data in the name of "correcting" it. Please consider these points:

It appears from the proposed rule that anyone would be able to make a charge that data or information are flawed --- incorrect, not objective, not useful, or without integrity --- and the burden of proof would be on the federal agency or data source. There should be a clear burden of proof on the person or group who make the initial charge. That proof should demonstrate expertise in the area of the data or information collected (e.g., statistical reanalysis or documentation of flawed procedure), not anecdotal information or a disagreement with the policy implications of the data. One must have a way to weed out frivolous charges based upon policy issues rather than the actual data. The federal agency would require staff expertise to adjudicate these charges.

2. What does it mean for a person to "seek and obtain corrections?" How can a person who did not collect the data, such as a complaining individual or a federal agency custodian of the data, know what remedy would be required? Data and information are the result of scientific processes of data collection. Would the agency be required to repeat experiments and field observations? ... repeat the original peer review process? One could not simply take the word of the person who contests the data.
3. Any post hoc "correction" of a database or scientific information flies in the face of data integrity. The very term "data integrity" means "wholeness" and "not tampered with" (by your own document). Individuals can present alternate

4. evidence/data or new interpretations, but cannot “correct” another person’s data any more than the original investigator can alter his or her own data. To do so is to commit a very serious ethical failure. Certainly the OMB does not want to encourage such actions.

5. Another interpretation of the word “integrity” refers not to the data, but to the individual who developed the data. But if one is objecting to the integrity of the person who developed or interpreted the data, there is still no way to “correct” the data other than to label the data set as having been challenged. A challenge would not mean that the challenger was correct in his/her charge against the data provider.

6. Any person who charges that data are flawed under this rule should be required to openly state his/her relationship with the data/information (familiarity and expertise) and provide information on his/her interest in it. For example, is the person a tobacco lobbyist objecting to data that shows smoking to be harmful? Is a group that opposes genetic engineering objecting to data that supports it? Such persons with viewpoints will make up the bulk of the commenters and that is fair, but comments should be made in context, just as scientists are required to state their funding sources and personal interests as a part of normal research and publication ethics.

The entities that are the focus of this rule --- Quality, Objectivity, Utility, and Integrity --- are themselves relative. Their interpretation is in the eye of the beholder. Who makes the decision as to the quality or utility of a piece of research? Utility is particularly dangerous to assess. We suspect that a person who objects to a particular governmental policy may well feel that data supporting that policy are not relevant. In order to make this rule enforceable, some hard criteria must be established. As for utility, one should allow a wide interpretation --- any data that can be argued to bear upon a particular problem, or that are used by scientists/citizens as context for interpretation of a problem or set of issues, have utility. Of course, basic research is in this second category. We are hard pressed to see a case where data acquired by federal agencies would lack utility. We worry that a narrow definition would be used by some individuals to try to eliminate information and thus weaken conclusions that they do not agree with.

8. We would oppose a definition of data quality tied to reproducibility. While this is often done in laboratory work, it is not possible to re-run a natural experiment or monitoring study such as is involved in environmental field research. We fear that use of “reproducibility” as a criterion might be used to discredit legitimate information derived from fieldwork and other historical datasets.

9. The requirement for federal agencies to report back to OMB on the number of complaints is troubling because it implies that an agency should strive for fewer complaints. In reality, the number of complaints is correlated not with the quality of data the agency handles, but with the level of controversy of the subject matter. If agencies are working to solve difficult social and scientific problems, they must collect data that put them in the line of fire. EPA and the Fish and Wildlife Service (endangered species act, etc.) are naturally going to be lightning rods. The rule should clearly state that OMB is aware that number of complaints will not be correlated with how good a job the agency is doing in meeting the mandates of the rule.

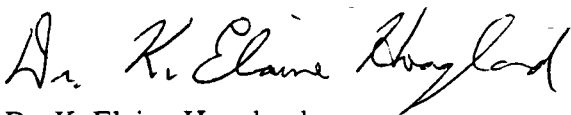
The answer to most of our concerns is stated in the draft rule itself: OMB is correct to say that GIVEN PROPER DISCLOSURE, the public and professional data users can assess for themselves the quality, objectivity, utility, and integrity of data/information collected by the federal government. The most straightforward way to deal with the task OMB has in writing this rule is to emphasize disclosure rather than “correction” or changes in data. “Correction” of data sets or information should be defined as adding contextual

information (better metadata), not as modifying or deleting the actual data. In the spirit of disclosure, information about the persons requesting the correction should be included as part of the amended metadata. All federal agencies should keep good metadata records, revealing the source, method of collection, dates of collection, and other related information. EPA and the Department of Interior have been working hard on metadata standards for at least ten years, and can be consulted on this issue.

If there is a problem with data quality in the federal government, combat it with more information, so that citizens in a free society can make choices based upon the full information available. We worry that the current draft rule could cause agencies to withhold or disregard data for fear of challenge, or worse, to react to special interests by altering and hence destroying the integrity of the information that has come to it through normal channels of review.

Thank you for the opportunity to comment.

Sincerely,

A handwritten signature in black ink that reads "Dr. K. Elaine Hoagland". The signature is written in a cursive style with a large, prominent "H" and "L".

Dr. K. Elaine Hoagland
National Executive Officer