DOCUMENT RESUME

ED 137 414 TM 006 223

AUTHOR Everett, Bruce E.
TITLE Confidentiality Methods in Educational Evaluations.
PUB DATE [Apr 77]
NOTE 10p.; Paper presented at the Annual Meeting of the American Educational Research Association (61st, New York, New York, April 4-8, 1977)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Administrative Agencies; Agency Role; *Civil Liberties; Computer Storage Devices; *Confidentiality; Data Bases; *Data Collection; *Educational Research; Evaluation Methods; Federal Aid; *Federal Legislation; *Guidelines; Legal Responsibility; Security
IDENTIFIERS Colorado

ABSTRACT

Current regulations and guidelines established by the Department of Health, Education and Welfare for the maintenance of confidentiality in educational research are summarized. Key terms, such as "system of records" are defined and elaborated upon. The responsibilities of the funding agency and its research contractors are enumerated and explained. Recommended procedures for ensuring compliance with the privacy and confidentiality statutes and regulations are discussed, especially regarding the actual collection of data and the maintenance of computerized data files. (Author/MV)

CONFIDENTIALITY METHODS IN EDUCATIONAL EVALUATIONS *

Bruce E. Everett

American Institutes for Research

At present, there are four major Federal confidentiality privacy statutes which regularly impinge upon educational evaluations. These are the Privacy Act, (PL93-579, 5USC 552a), the Protection of Human Subjects clause in the National Research Act (PL93-348, 42USC, 2891-1), the Family Educational Rights and Privacy Act (the "Buckley Amendment") (PL93-380, Section 438, 20USC 1232g), and the Freedom of Information Act (PL93-502, 5 USC 552). Inasmuch as a summary of each statute is readily available (Weinberger and Michael, 1976; 1977; Michael and Weinberger, 1977), neither their detailed regulations nor their redundancies and noneducational clauses will be discussed here. Rather, this paper will concentrate on the guidelines and procedures that are currently necessary for the general compliance with these statutes as far as educational researchers are concerned.

Generally speaking, these statutes place the responsibility for compliant educational research on both the Federal funding agency (i.e., DHEW) as well as the actual researcher. In order to comply with these statutes, the researcher must first be able to describe three things about his or her data: 1) does it include information that is personally identifiable; 2) If so, who has access to that information; 3) Has the individual been properly informed of his rights and of the consequences regarding his participation in the research (If the individual is a minor, has the parent or guardian been properly informed?)?

If the data collected is not personally identifiable, then most of the privacy regulations become moot. The one exception is that the individual must be informed of this complete anonymity. Also, the intention to maintain anonymity is not the same as actually doing so from the start: as long as the potential for individual identification exists, the researcher has definite legal responsibilities to ensure confidentiality .

Here are certain key terms that should be kept in mind:

*A presentation given at the 1977 annual meeting of the American Educational Research Association, ~~[illegible]~~ N.Y., N.Y.

• human subjects - not merely limited to people on whom experiments are performed. Current interpretation of the law also places questionnaire respondents, interviewers and observers in this category.

• individual - a living citizen of the United States or an alien lawfully admitted for permanent residence. In the case of minors, parents or guardians are also included. A business firm which is identified by the name of one or more individuals is not considered as an individual under the Privacy Act.

• maintain - to maintain, operate, collect, use or disseminate.

• record - any item, collection or grouping of information about an individual by an agency or its contractors maintained for a specified purpose that contains his or her name, identifying number, symbol, fingerprints, voiceprints, photographs or any other means of direct identification. "Identifying number" refers to Social Security numbers, drivers' licenses, draft registration numbers, etc., which are clearly and unambiguously linked to individuals. Arbitrary ID numbers, so long as they cannot possibly be translated into these direct forms of personal identification, do not count.

• system of records - a group of any records under the control of any agency or its contractors from which data can be retrieved by the name of the individual or any other identifying particular assigned to that individual.

• agency - for the purposes of privacy legislation, the source of funding, i.e., the Department of Health, Education and Welfare.

• privacy - the right of an individual to 1) determine what records pertaining to him or her are being maintained by Federal agencies, 2) prevent such records obtained for a particular purpose from being used for some other purpose without his or her consent, 3) gain access to, copy, correct, or amend such records, 4) be assured that such records have been maintained for lawful purposes and 5) have redress if the maintenance of such records is deemed unlawful.

• confidentiality - the obligation of an agency and its contractors to ensure the privacy rights of the individuals on whom they maintain records.

• security - specific methods to ensure confidentiality, especially the prevention of unauthorized access.

• consent - the authorization, written or otherwise, by an individual for an agency and its contractors to maintain a record pertaining to that

individual. Current practice allows for positive consent to be assumed if individuals have been informed in advance of the creation of a system of records and there is the opportunity for written refusal for consent.

- access - the ability to acquire, directly or indirectly, the information contained in a system of records.
- authorization for disclosure - restricted to the system of records manager within the sponsoring agency, i.e., the Project Officer.

In order to streamline the process of establishing compliance with these statutes, the various Federal agencies within DHEW have prepared guidelines for compliance as well as official reporting forms. The best known of these deal with the Protection of Human Subjects, and require that proposed research procedures be reviewed and certified by an independent panel prior to actual work on a contract or grant. Because the improper release of personal data can be harmful to a respondent, a description of the confidentiality procedures in force and of the planned uses of the data has become a necessary input to the panel's decision.

Another form which has come into use is a statement of compliance with the Privacy Act. Although it is signed by a representative of the sponsoring agency, the researcher typically has to provide the actual information. Several key questions about the data must be answered. First, does it qualify as a "system of records", that is, is the data personally identifiable? Second, is there "routine use" of the data as a system of records, that is, are individuals constantly and systematically being personally identified? Third, who "maintains" the data and is in control of access to it?

In addition to certification of compliance with privacy and confidentiality regulations, agencies and their contractors are directed to follow specific guidelines in order to ensure compliance. They are currently required to:

- establish administrative, technical and physical safeguards in accordance with DHEW and National Bureau of Standards guidelines to ensure the security and confidentiality of a system of records;
- establish rules of conduct for all employees involved in the design, development, operation or maintenance of a system of records and inform them of the Federal rules and regulations and penalties for noncompliance that apply;

- report to the system of records manager all requests for disclosure of an individual's record, requests by any individual for access to his or her record and requests by any individual to amend his or her record
- refrain from any unauthorized disclosure of a record to anyone outside the sponsoring agency
- provide for the authorized access to or amendment of individual records by the individuals themselves
- keep an accurate accounting of all disclosures
- make such additional reports regarding the maintenance and operation of the system of records as required by contract or by law

Contractors also have additional obligations before creating a system of records, namely, to:

- inform the system of records manager within the sponsoring agency, i.e., their Project Officer or monitor, of all the information necessary to describe the existence and character of the proposed system of records pursuant to publication of that description in the Federal Register
- refrain from maintaining such a system of records for thirty days after the publication of the description of the proposed system in the Federal Register
- maintain in the system of records only that information which is relevant and necessary to the purposes of the contract
- collect information directly from the individual to the greatest extent possible, e.g., avoid hearsay and second-hand sources
- inform each individual from whom information is requested by means of a direct statement on the form used to collect information or on a separate form that the individual can retain of a) the principal purpose or purposes for which the information will be used, including specific enumeration of the purpose of the contract, b) the routine uses, i.e., disclosure of personally identifiable information, of the system of records and c) the effects on the individual, if any, of not providing any or all of the requested information

More specific guidelines for the ensurance of privacy and confidentiality include the following:

- all personnel having access to or responsibility for a system of records are required to take and sign a nondisclosure oath and to be informed

of their responsibilities and obligations regarding confidentiality

- all records and associated documents are to be stored in a locked receptacle when not in use
- all records and associated documents are to be inventoried and accounted for
- data banks and files shall be protected by passwords and other techniques which can be used to identify and verify the system user
- only personnel authorized to have physical or online terminal access to a system of records are allowed to do so
- backup and recovery data sets are subject to the same security requirements as primary data sets
- direct use of personally identifiable data should be marked "FOR OFFICIAL USE ONLY"
- a system of inspection and monitoring should be in effect to ensure that security measures are being properly adhered to
- all instances of access to a system of records should be authorized in advance and logged
- all anomalies in access to the systems of records should be thoroughly investigated and described
- disposal and destruction of unneeded records should include security measures; e.g., shredding before disposal

The problem of maintaining confidentiality begins with the actual collection of data. As soon as an individual becomes identified with a given research effort or information about that individual enters a system of records, the confidentiality process must be operative. One must keep in mind that any information, even that which is not harmful to the respondent, such as age and height, constitutes part of a system of records if it is personally identifiable. Admittedly, the stringency of the security measures used to safeguard such information may not have to be as great as when obviously harmful information such as criminal records are part of a system of records, but the authority for easing security measures lies with the system of records manager and not with the researcher.

The most direct method to establish information about individuals as a system of records is to have their names placed directly on the data collection forms. As long as those names remain on the forms, they constitute a

continuous security problem. For that reason, the direct identification method may not be allowed by the sponsoring agency. One compromise to this method is to have the respondent name located on a perforated "tear-away" portion of the data collection form: once data processing is complete, the names can be torn off the forms, thereby making the forms physically anonymous. Another variation of the "tear-away ID" method is to use pressure-sensitive stick-on labels that have two parts, one with the respondent's name on it that can be peeled off and the other with an arbitrary ID number on it that remains fixed to the data collection form. Still another version is to have the name of the respondent on the outside of an envelope and the data collection form with an ID number on it enclosed in the envelope.

Implicit in these methods is the knowledge on the part of the researcher of just who the respondents are. If the system of records manager decides that this knowledge is a liability to the respondents, more elaborate methods of data collection are needed. In the "double blind" method, someone other than the researcher contacts the respondents, collects their data and delivers it to the researcher. This outside entity acts as an "escrow agent" for the information linking an individual to his or her data. A less stringent but more manageable method is to follow the same procedures in-house while maintaining the data base. By keeping personal identification information in a separate but cross-linked file from actual research data, the research data is not in and of itself a system of records. It only becomes one when it is merged with the personal identification file.

In small studies, the maintenance of acceptable confidentiality measures may be quite simple, e.g., by locking up all data in a safe when not being used, and by working with the data in a locked room. In short, by limiting access to data to a limited number of individuals under secure conditions, one has control over the confidentiality of the data.

If the system of records is stored in a computer system, the problems of maintaining confidentiality are increased. Physical access, other than the theft of card decks or magnetic tapes, is less of a problem than symbolic access, that is, the ability to use valid computer-related operations to inspect, copy, manipulate or destroy someone else's system of records.

In most large computer environments, there are hundreds, even thousands

of users, each of whom has a perfect right and reason to use the computer. In order to gain unauthorized access to a system of records, the outsider has to know 1) that the system exists, 2) how and where it is stored, 3) what name it is stored under and 4) what passwords and other protection devices have been used to restrict access. The assumption that this information is only of interest to authorized users and is known only by them is extremely naive.

Take, for example, the choice of keywords and passwords. In order not to forget their own keywords, many users will use their own initials or a convenient acronym (TEP for Teacher Evaluation Project). Access to such "protected" data is available to anyone with a casual knowledge of the study. A more devious but not difficult method is to appropriate discarded computer listings and coding sheets, since keywords are typically disguised by the computer on output, not input. Another common mistake is to write down keywords and openly display them or let others share them.

Here are some simple rules for maintaining keyword protected computer files:

- have an explicit list of authorized users on file with the computer center.
- choose the letter or digit of each keyword at random and change it often. This may be inconvenient if there are multiple users, but the protection is worth it.
- restrict the number of occasions the keyword is written down to a minimum. Posting keywords on a bulletin board or circulating them in memos defeats their purpose.
- try to control the disposal of computer cards and paper that contains keywords. The overstrike feature that many on-line terminal systems use can be read through quite easily.
- unauthorized access, once suspected, can be monitored by the computer system itself. A "spy" circuit can be used to signal and locate the source of each access.
- the naming of data sets can also have confidentiality implications. Many computer installations publish daily the names of the on-line data sets that each user has, and there is a vast difference in the amount of curiosity aroused by a data set called "Data" and one called "Teacher

Salary Data."

The efforts within DHEW to establish guidelines in defining privacy and confidentiality compliance are still at an interim stage and are likely to change as new legislation and new research concerns arise. In particular, the inherent contradiction between maintaining data confidentiality and allowing access to information about an individual create real problems for the researcher. If data are not collected in a personally identifiable manner, how can the researcher authenticate the responses? Moreover, how can that data be reliably linked to other individual level data from another source? The collection of completely anonymous data rules out the possibility for later verification of responses, longitudinal research and multi-instrument designs. Also, the typical requirement that raw data be held for several years after the end of a research contract creates a situation where the legal obligations and liabilities of the contractor outlast the financial coverage provided by the contracting agency.

As new legislation is enacted and the interaction between sponsor and researcher exposes the advantages and disadvantages of current guidelines and regulations, the nature of the confidentiality process is bound to change. Nevertheless, the maintenance of the confidentiality of research data is now a legal obligation on the part of the researcher and not simply a professional or ethical one. By building confidentiality methods directly into the evaluational design, they may prove to be less of a hindrance than one might expect. If, however, they are postponed or ignored, the researcher is definitely taking severe risks through noncompliance with the law.

BIBLIOGRAPHY

Davis, C. R. The Buckley regulations: Rights and restraints. Educational Researcher, February 1975, 4(2)

Everett, B. E. Effective data management and quality control techniques for large-scale longitudinal research. A paper presented at the 1975 annual meeting of the American Educational Research Association, Washington, D.C.

Federal Property Regulations, temporary Regulation E-34. Interim adequate measures to safeguard "Personal Information".

Institutional guide to DHEW policy on protection of human subjects. DHEW Publication No. 72-102. December, 1971.

Martin, J. Security accuracy and privacy in computer systems. Englewood Cliffs, N.J.: Prentiss-Hall, Inc., 1973.

Michael, J. A. and Weinberger,J. A. Federal restrictions on educational research: Protection for research participants. Educational Researcher, January, 1977, 6(1).

National Bureau of Standards. Computer security guidelines for implementing the privacy act of 1974.

National Institutes for Health. Privacy act of 1974 announcement. NIH Guide for grants and contracts. 5(5), April 28, 1976.

Office of Management and Budget. Privacy act guidelines. Federal Register, July 9, 1975, 28948-28978.

Weinberger, J. A. and Michael, J. A. Federal restrictions on educational research. Educational Researcher, December 1976, 5(11).

Weinberger, J. A. and Michael, J. A. Federal restrictions on educational research: A status report on the privacy act. Educational Researcher, February 1977, 6(2).