DOCUMENT RESUME

ED 416 243 TM 028 123

AUTHOR	Ahmed, Susan; Barbett, Samuel; Carr, Peggy; Frase, Mary; Hoffman, Lee; Kasprzyk, Daniel; Kolstad, Andrew; Owings, Jeffrey; Snyder, Tom
TITLE	Improving Data Quality in NCES: Database-to-Report Process. Working Papers Series.
INSTITUTION	National Center for Education Statistics (ED), Washington, DC.
REPORT NO	NCES-WP-97-13
PUB DATE	1997-04-00
NOTE	78p.; Prepared by the Task Force on Quality Systems.
AVAILABLE FROM	U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, 555 New Jersey Avenue, N.W., Room 400, Washington, DC 20208-5654.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	Academic Achievement; Data Collection; *Databases; Educational Trends; Elementary Secondary Education; *National Surveys; Postsecondary Education; *Quality Control; Research Methodology; *Research Reports
IDENTIFIERS	*Data Quality; Digest of Education Statistics; National Assessment of Educational Progress; *National Center for Education Statistics; Schools and Staffing Survey (NCES)

ABSTRACT

The Task Force on Quality Control of the National Center for Education Statistics (NCES) was created to study the quality control procedures used for NCES publications and to determine if there are ways to improve the current procedures. Because a detailed review of the entire survey process would have been too time consuming, the Task Force decided to focus on errors that occur after the survey datafile has been accepted. To study the existing quality control procedures, the Task Force studied the database-to-report processes of the: (1) National Assessment of Educational Progress 1994 Trends in Academic Progress; (2) Digest of Education Statistics; (3) Schools and Staffing Survey in the United States: A Statistical Profile; and (4) Fall Staff in Postsecondary Institutions, 1993. Interviews were conducted with NCES staff and representatives of other government agencies. Overall, the Task Force found that there were no major systematic problems affecting the data being released by the NCES. There are quality control procedures in place at strategic points, and the general situation is very positive. Some possible areas for improvement were noted, and a list of recommendations, ultimately reduced to five, was prepared to improve the quality of NCES products. Typically, problems could occur in data compilation, documentation, formatting, and structural areas. Appendixes contain a memo on quality control practices, a discussion of the application of recommendations to previously identified problems, and references. (Contains one figure and seven references.) (SLD)

ED 416 243

NATIONAL CENTER FOR EDUCATION STATISTICS

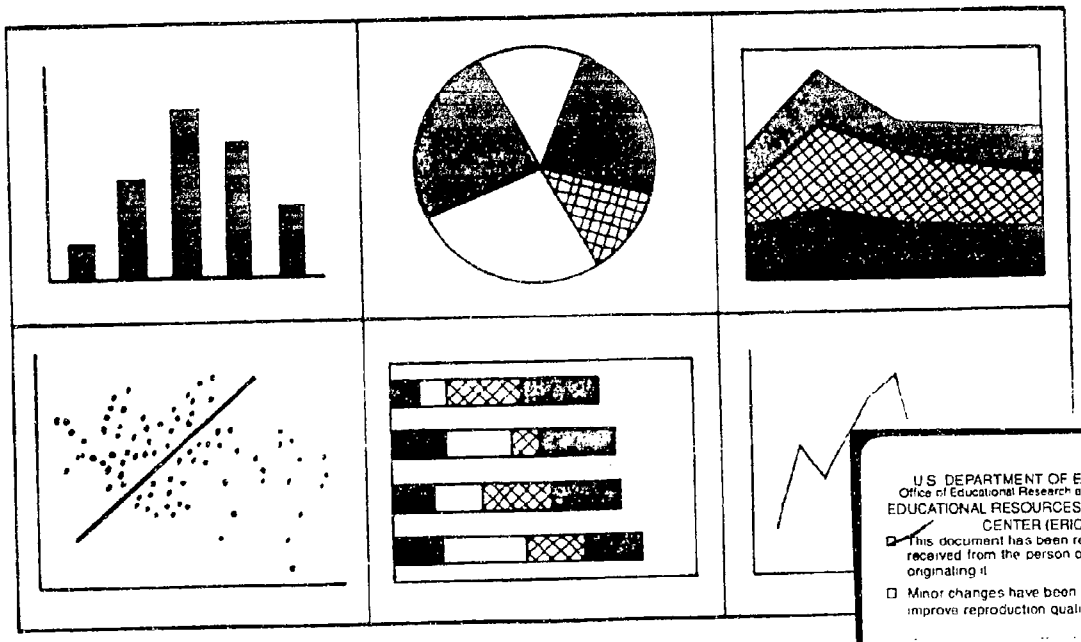
Working Paper Series

Improving Data Quality in NCES:

Database-to-Report Process

Working Paper No. 97-13

April 1997



U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

U.S. Department of Education
Office of Educational Research and Improvement

Improving Data Quality in NCES:

Database-to-Report Process

Working Paper No. 97-13 April 1997

Contact: Susan W. Ahmed
Statistical Standards and Services Group
(202) 219-1781
e-mail: susan_ahmed@ed.gov

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
Ramon C. Cortines
Acting Assistant Secretary

National Center for Education Statistics
Pascal D. Forgione, Jr.
Commissioner

Statistical Standards and Services Group
Susan W. Ahmed
Chief Mathematical Statistician

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Office of Educational Research and Improvement
U.S. Department of Education
555 New Jersey Avenue, NW
Washington, DC 20208

Suggested Citation

U.S. Department of Education. National Center for Education Statistics. *Improving Data Quality in NCES: Database-to-Report Process*, Working Paper No. 97-13, by Task Force on Quality Systems. Washington, D.C.: 1997.

April 1997

Foreword

Each year a large number of written documents are generated by NCES staff and individuals commissioned by NCES which provide preliminary analyses of survey results and address technical, methodological, and evaluation issues. Even though they are not formally published, these documents reflect a tremendous amount of unique expertise, knowledge, and experience.

The ***Working Paper Series*** was created in order to preserve the information contained in these documents and to promote the sharing of valuable work experience and knowledge. However, these documents were prepared under different formats and did not undergo vigorous NCES publication review and editing prior to their inclusion in the series. Consequently, we encourage users of the series to consult the individual authors for citations.

To receive information about submitting manuscripts or obtaining copies of the series, please contact Ruth R. Harris at (202) 219-1831 or U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, 555 New Jersey Ave., N.W., Room 400, Washington, D.C. 20208-5654.

Susan Ahmed
Chief Mathematical Statistician
Statistical Standards and
Services Group

Samuel S. Peng
Director
Methodology, Training, and
Service Program

Improving Data Quality in NCES:

Database-to-Report Process

Prepared by:

Task Force on Quality Systems

Members:

National Center for Education Statistics
Susan Ahmed, SSSG, Chair
Samuel Barbett, SCSG
Peggy Carr, AG
Mary Frase, DDLSG
Lee Hoffman, SCSG
Daniel Kasprzyk, SCSG
Andrew Kolstad, AG
Jeffrey Owings, DDLSG
Tom Snyder, DDLSG

Support Staff:

Education Statistics Services Institute
Karol Krotki
Matthew Riggan

Westat
David Marker

April 1997

Table of Contents

EXECUTIVE SUMMARY 1

INTRODUCTION 1
METHODOLOGY 1
RESULTS OF THE STUDY 2
Typical problems in the database-to-report process 3
Quality control practices at NCES 4
RECOMMENDATIONS 6
Themes of Recommendations 6
Final Recommendations 7

I. ANALYSIS OF RESULTS 9

BACKGROUND INFORMATION 9
Focus of the Study 10
Literature on Quality Control in the Last Stages of the Survey Process 12
SECTION 1. ANALYSIS OF THE DATABASE-TO-REPORT PROCESS 14
Identifying steps in the database-to-report process 14
The database-to-report process of four reports 18
SECTION 2. TYPICAL PROBLEMS IN THE DATABASE-TO-REPORT PROCESS 22
2.1 Data Compilation Errors 22
2.2 Inadequate Documentation 23
2.3 Table Formatting Errors 25
2.4 Organizational Characteristics 26
SECTION 3. QUALITY CONTROL PRACTICES AT NCES 28
3.1 Phase 1: Early Review Procedures 29
3.2 Phase 2: Group Review and Adjudication 31
3.3 Post-Adjudication Review 33
3.4 Additional Quality Control Measures 34
SECTION 4. PRODUCTION OF REPORTS AT OTHER FEDERAL STATISTICAL AGENCIES 36
4.1 Bureau of the Census: Statistical Abstract 36
4.2 Bureau of the Census: Demographic Statistical Methods Division 37
4.3 National Center for Health Statistics 38
4.4 Bureau of Labor Statistics: Publications 38
4.5 Bureau of Labor Statistics: Current Price Index 39
4.6 National Agricultural Statistics Service 39
4.7 Statistics Canada 40

II. RECOMMENDATIONS BASED ON FINDINGS 41

SECTION 1. THEMES OF THE RECOMMENDATIONS 41
SECTION 2. FINDINGS AND PRELIMINARY RECOMMENDATIONS 43
SECTION 3. FINAL RECOMMENDATIONS 51
1. Develop a handbook of technical guidelines for NCES publications 51
2. Include input from survey staff 52
3. Improve communication between NCES staff and programmers 52
4. Formalize internal documentation of tables 53
5. Increase NCES staff available for review and increase use of contractor help in the review process 54
Management Participation 55
SECTION 4. FUTURE WORK 56

APPENDICES 57

APPENDIX A. MEMO ON NEW QUALITY CONTROL PROCEDURES 57
APPENDIX B. APPLICATION OF RECOMMENDATIONS TO PREVIOUSLY IDENTIFIED PROBLEMS 64
APPENDIX C. REFERENCES 65

EXECUTIVE SUMMARY

Introduction

The National Center for Education Statistics (NCES) Task Force on Quality Systems was created to study the quality control procedures used for NCES publications and to determine if there are ways to improve the current procedures in a cost effective manner. Although the Task Force was originally created as the result of a few errors that were found in some NCES publications (see Appendix B), it was charged with the broader task of looking at the general issue of data quality in all NCES publications.

Initially, the Task Force began examining the entire survey process, from planning stages through the final report production. However, because a detailed review of the entire process would have been too time consuming and because current quality control measures in early stages of the survey process have not revealed notable problems, the Task Force decided to focus on errors that occur after the survey datafile has been accepted. The Task Force not only examined procedures within NCES, but also the practices of NCES contractors and several Federal statistical agencies.

Specifically, this report examines:

- the database-to-report process for NCES publications,
- typical problems encountered in the process,
- quality control practices at NCES, and
- production of reports at other Federal statistical agencies.

After these sections, findings and recommendations are discussed, and future work is identified. Readers solely interested in the recommendations of the Task Force may proceed to Chapter II.

Methodology

To better identify the existing quality control procedures used at NCES, the Task Force studied the database-to-report processes of four reports: *National Assessment of Educational Progress 1994 Trends in Academic Progress; Digest of Education Statistics; Schools and Staffing Survey in the United States: A Statistical Profile, 1993-1994;* and *Fall Staff in*

BEST COPY AVAILABLE

Postsecondary Institutions, 1993. We interviewed NCES staff, as well as contractors involved with the reports, about current quality control procedures. In addition, we organized a discussion group with several individuals involved in the production of *Digest of Education Statistics,* and an individual responsible for the Common Core of Data (CCD) reports was interviewed. An interview was also held with staff from the Bureau of the Census to examine the quality control procedures used for *Statistical Abstract of the U.S.* Finally, we contacted individuals from several other statistical agencies about quality control practices. The agencies contacted were The National Center for Health Statistics, the Bureau of Labor Statistics, the National Agricultural Statistics Service, and Statistics Canada. In addition to the interviews, a literature review was done on quality control in the survey process.

The Task Force developed two flowcharts that demonstrate the survey process. Figure 1 outlines the steps taken from the planning stages all the way through publication. The second flowchart (Figure 2) further breaks down the last three steps of table production, report drafting, and final publication. This second chart was used as a basis for developing the interview questions and format. Based on the information gained in the interviews, the Task Force also created flowcharts detailing the final review stages for each NCES report. While the interviews provided information on the similarities and differences between the database-to-report processes of each publication, prior NCES documentation as well as input from Task Force members supplemented the interviews to help in outlining the review and adjudication process.

Based on the information gathered in its meetings and through interviews, the Task Force compiled an original list of 21 recommendations, which varied greatly in scope. Upon further review, five final recommendations were selected for initial action and arguments were developed for why these recommendations should be implemented. In addition, the Task Force considered suggestions for how NCES might begin to implement the final recommendations.

Results of the Study

Overall, the Task Force found that there were no major systematic problems affecting data being released by NCES. There are quality control procedures in place at strategic points and, whereas some possible areas for improvement were identified, the general situation is very positive.

The literature review on quality control and error reduction indicated that little has been written about the later stages of the survey process, which are the focus of this study. Although the literature notes that consistency in treatment of data is imperative, few recommendations are made for preventing errors resulting from inconsistent data.

The development of the second flowchart of the database-to-report process provided a breakdown of the general procedures currently in place at NCES. The database-to-report process involves the following sequence of steps:

1. Preliminary table generation and an internal review by the contractor or NCES Project Officer

2. Table formatting and report writing to accompany the tables

3. Second review phase, consisting of several reviews occurring concurrently or sequentially

4. Adjudication

Typical problems in the database-to-report process

The Task Force was able to identify and analyze typical errors in the database-to-report process. Errors tend to occur most frequently in three major stages of report development: production of preliminary tables, formatting and documentation of tables, and post-adjudication revisions.

Four general types of errors may occur in the database-to-report process:

- Data compilations errors
- Documentation errors
- Formatting errors
- Structural sources of error

Data compilation errors occur when data are processed through NCES or the contractor. These errors are likely because data come from a variety of sources and in a variety of print and electronic formats. Another cause of data compilation errors is the change in survey variable definitions over time, which can lead to misinterpretation of the meaning of the data.

The second type of error, documentation errors, result from inadequate information about the data included in a table or file, generally due to a lack of documentation about where the data came from or how they have been processed. When specific instructions are not given to the contractor on the exact variables to use, they may be selected at the discretion of the programmer. Also, the definitions for data labels may vary, leading to inconsistent treatment of data. Depending on how terms are defined, different values could appear for each label, or data could be processed differently than the way they are reported. Another type of documentation error is the inadequate documentation of null data.

Formatting errors are a third type of error in the publication process. They can occur when there is a change in the format of the table. For example, formatting changes may occasionally shift columns but not data labels.

Structural sources of error result from a lack of established channels of communication between NCES Groups working on different publications. For example, one Group may not be aware of another Group's publication which might be treating the same data differently and publicizing results which are at odds with those of the first Group. Even within a Group, simultaneous work can take place on different versions of a draft. Some publications have guidelines for maintaining a "master copy" of documents, although this procedure is not standardized throughout NCES. Finally, administrative pressure can contribute to error. For example, the publication format may be altered, based on changing data importance, or reports may need to be produced quickly, due to immediate policy-relevance. All of these factors may lead to an increased number of errors in the publication process.

Quality control practices at NCES

Extensive quality control checks have been developed by NCES for all reports and tables. The reviews of NCES reports include cross-checking tables against original tables, other reports that use the same data, or reports from previous years.

In the early review phase, the first focus is on the review of preliminary tables. Next is a review of the completed draft before it enters the full NCES review. The table reviews are conducted primarily by the contractor, or in some cases, the NCES staff in the program area. The second part of the early review process is a Program and Group review, which assures that the report is ready for peer review. This stage of the process

involves a comprehensive and time-consuming review of the draft for approval at the Group level. At the end of the Group review, the Senior Technical Advisor signs off on most reports prior to submission for adjudication.

The next review phase is uniform for all NCES publications. Peer reviewers from inside and outside the agency receive the draft and prepare comments. These comments are discussed with the author at the adjudication meeting. Once all revisions have been incorporated, a memo is created to highlight the changes to the document. After adjudication, the report is signed off by the adjudicator and the Chief Statistician.

For certain publications, additional quality control strategies have been implemented. These include

- identifying tables and sections which are error-prone and listing the ways to check such sections to reduce error,
- documenting all changes made to reports at each review stage, and
- automating the report process at certain points.

Many of these strategies are currently used in NCES and other Federal statistical agencies to reduce errors in publications.

BEST COPY AVAILABLE

Recommendations

After a preliminary list of recommendations was developed, the Task Force categorized them into five general themes which had emerged throughout their development. Each of the recommendations to follow will be discussed in greater detail in Chapter II of this report, "Findings and Recommendations." In addition, five final recommendations were selected for immediate action from those listed below. These final recommendations, detailed at the end of Chapter II, are highlighted in boldface here and listed separately below.

Themes of Recommendations

Spread Existing Skills Within NCES

- ***Develop a handbook of technical guidelines for NCES publications***
- Train appropriate NCES staff in statistical software applications

Improve Documentation and Communication

- Front end
  - ***Include input from survey staff***
  - ***Improve communication between NCES staff and programmers***
  - Improve documentation of data used to create tables
  - Improve documentation of database status
  - Improve consistency across publications*
  - Improve review criteria for checking tables*
- Back end
  - ***Formalize internal documentation of tables****
  - Improve archiving of NCES survey datafiles

Improve Uses of Computer Systems for Checking Key Results

- Generate supplementary cross tabulations*
- Develop master analysis files for a report*
- Provide additional programming support for the Chief Statistician

Improve the Review Process

- ***Increase use of contractor help in the review process***
- ***Increase NCES staff available for review***

- Define objectives and responsibilities for each review step
- Delineate clearly activities, responsibilities, and timing for each review step
- Share specialized review procedures throughout NCES

Improve Contract Procedures

- Allocate specific funds for data quality in contracts*
- Improve the use of data analysis plans*

* These recommendations could be addressed specifically by some programs rather than being implemented universally throughout NCES. They represent changes that could be implemented immediately and differentially across the agency.

Final Recommendations

The Task Force selected five recommendations for initial implementation. The selection criteria were cost, feasibility, and potential benefits.

1. Develop a handbook of technical guidelines for NCES publications
2. Include input from survey staff
3. Improve communication between NCES staff and programmers
4. Formalize internal documentation of tables
5. Increase NCES staff available for review and increase use of contractor help in the review process

I. ANALYSIS OF RESULTS

Background Information

NCES Commissioner Pascal D. Forgione, Jr., created the National Center for Education Statistics Task Force on Quality Systems in September 1996, to study the quality control procedures for NCES publications and determine if there are ways to improve the current procedures in a cost effective manner. Chaired by NCES Chief Statistician Susan Ahmed, the Task Force consists of nine representatives from NCES and is supported by one representative from Westat and two from the Education Statistics Services Institute (ESSI)[1]. Although the Task Force was created because four errors had occurred in recent NCES publications (see Appendix B), its investigations are not limited to the areas in which those errors occurred. Instead, the Task Force is charged with taking a fresh look at the issue of ensuring data quality in NCES publications.

One of the first attempts to look at this issue, an attempt that predates the creation of this Task Force, is presented in Appendix A. In fact, some of the work done by the Task Force on Quality Systems used this document as a starting point and early foundation for this report.

The report of the Task Force consists of an Executive Summary, Chapter I, "Analysis of Results", and Chapter II, "Recommendations Based on the Findings". Chapter I begins with a "Focus of the Study" and a literature review. The remainder of Chapter I, divided into four sections, provides background information and context which serves as the basis for the recommendations detailed in Chapter II. The four sections of Chapter I cover the following topics:

- Section 1: Beginning with a description of the steps that comprise the late stages of the database-to-report process, "Analysis of the Database-to-Report Process" analyzes the ways these steps fit into the overall process and examines the variations in these steps for different NCES publications.

[1] The NCES members are Susan Ahmed (Chair), Samuel Barbett, Peggy Carr, Mary Frase, Lee Hoffman, Daniel Kasprzyk, Andrew Kolstad, Jeffrey Owings, and Tom Snyder. Support staff included David Marker (Westat), and Karol Krotki and Matthew Riggan (ESSI).

- Section 2: "Typical Problems in the Database-to-Report Process" explores the types of problems that occur in these steps, providing specific examples of different kinds of error and exploring organizational characteristics that may contribute to the occurrence of error.

- Section 3: "Quality Control Practices at NCES" describes the measures currently in place to reduce error and maintain quality of NCES publications, focusing on specific practices and organizational characteristics that help to reduce error.

- Section 4: "Production of Reports at Other Federal Statistical Agencies" describes the review and quality control procedures for other Federal statistical agencies.

Chapter II presents findings and recommendations. Appendix B shows how the recommendations in Chapter II could be applied to the problems that originally motivated the creation of the Task Force.

Focus of the Study

In early meetings, the Task Force made basic decisions about the scope and breadth of its work. In looking at the entire survey process from the earliest stages of preparation and data collection through final publication, it determined that a full study of the whole process would be too complex and time consuming. Furthermore, there was a consensus that quality control measures in the early stages of this process are generally of acceptable quality. Therefore, the Task Force chose to focus on those problems that occur after the following activities have been completed:

1. A quality control process was implemented during the survey process.
2. A data review was completed during the survey process.
3. The data on the file are determined to be acceptable.

In deciding to focus on the steps of the survey process following the preparation of a clean datafile, the Task Force recognized that "clean" does not necessarily mean "perfect." Datafiles are edited to certain specifications, and perfect datafiles may not exist. For example, files are edited to jointly optimize the goals of data accuracy, timeliness, and cost. Consequently, in the interests of timeliness and cost, it might

happen that some observations are left unchanged, even though they may appear anomalous. This balancing of accuracy, timeliness, and cost is one of the perspectives in which to view the findings reported in this document.

The Task Force also decided to examine practices not only within NCES, but also for NCES contractors as well as for other Federal statistical agencies. This decision enabled the Task Force to examine other agencies for consistency of operations across programs as well as best practices in the field.

In order to better understand the existing quality control procedures at NCES, the Task Force studied four reports: *National Assessment of Educational Progress 1994 Trends in Academic Progress; Digest of Education Statistics; Schools and Staffing Survey in the United States: A Statistical Profile, 1993-1994;* and *Fall Staff in Postsecondary Institutions, 1993.* An attempt was made to select reports from different areas within NCES in such a way as to be representative of all major reports produced by NCES.

To identify the quality control procedures used to produce these reports, the Task Force interviewed individuals associated with these reports both at NCES and its contractors. In addition, a discussion group with several individuals associated with the *Digest of Education Statistics* specifically examined organizational and communication issues for staff working on a single recurring publication.

Finally, additional interviews were conducted with representatives from other statistical agencies. In one case the Task Force focused on a specific report (*Statistical Abstract* produced by the U.S. Bureau of the Census), while in others general quality control practices were discussed. In total, quality control practices were reviewed for five agencies outside of NCES: Bureau of the Census (BoC), National Center for Health Statistics (NCHS), Bureau of Labor Statistics (BLS), National Agriculture Statistics Service (NASS), and Statistics Canada.

The interviews were conducted by ESSI staff, with support on the discussion group from Westat. In designing and testing protocols, one additional interview was held with an individual responsible for the Common Core of Data (CCD) reports at NCES, which ultimately contributed to the final report. The material to be covered in interviews was determined through a number of Task Force meetings in which different types of problems and practices were discussed.

BEST COPY AVAILABLE

Literature on Quality Control in the Last Stages of the Survey Process

The vast majority of the literature dealing with quality control and error reduction procedures in statistics focuses on the actual error reduction in the survey process and the collection and imputation of data. Much of the literature discusses sophisticated techniques that are used to "clean" datasets, and significant steps that have been taken to reduce error in the collection process. Furthermore, much of the literature on data "quality," particularly the quality of Federal statistics, actually defines "quality" not by the presence or absence of error but rather by the utility of the data.

The work that does examine quality, and specifically error reduction, in the late stages of the production process tends to do so in passing within the larger context of general research methodology. Such information is generally found in handbooks and overall methodology guides. These works provide a front-to-back model for the design and execution of research projects. Their focus is not specifically on error reduction, but they do contain some useful observations in studying error in the production process.

For instance, Booth, Colomb, and Williams (1995) discuss changes in format that occur in the production of tables. When data are being moved and reconfigured, even though no actual calculations are being performed, there is a chance that error will occur. Aside from noting that consistency is imperative, however, research guides do not recommend any procedures for preventing such error.

Some research does focus on common errors in calculation. This work generally examines different types of mistakes made in analyzing the data from a "clean" dataset. In some cases they are mistakes of interpretation, while in others mistakes are caused by misapplication of tests or failure to report the data correctly (Wollins, 1982). Once again, this work examines the types of error that occur, but does not make any recommendations about how to systematically reduce such error.

While there may be a shortage of literature on specific types of quality control measures in the late stages of the production of statistical tables and reports, literature does exist which raises some general issues concerning data quality that are relevant to the scope of this investigation. The first is the use and inclusion of metadata in statistical databases. Within statistical agencies that are decentralized, there is always a risk that data without proper information regarding the source will be misused or misinterpreted, or that data products will appear incongruous or incompatible with other data from the same database. This potential for misuse of data greatly increases

when databases become public-use datasets and are accessible to large numbers of users. The rapid expansion of Internet technology further increases this potential. Sundgren and Dean (1996) argue that all data should be documented thoroughly enough so that

> future generations of researchers, living maybe 200 years from now, and who would have no first-hand knowledge of the data collections as such or the society in which they were collected, would be able to interpret and analyze the archived data. (Sundgren and Dean, 16.)

The authors go on to describe the documentation system of Statistics Sweden, which includes not only technical information but also statistical methodology and subject matter aspects. This type of documentation requires coordination among agencies to define key concepts used to describe statistical surveys. The goal was to make statistics user-friendly and accessible while retaining the autonomy of individual statistical agencies.

Yet, what exactly is quality? Inquiries into the quality of Federal statistics in the United States have examined this definition. In general, quality is evaluated in two ways: accuracy and timeliness. In a statement submitted to the Joint Economic Committee of the U.S. Congress, Michael Boskin, Chairman of the Council of Economic Advisors, writes

> In still other cases, the problem may be the familiar one of balancing the tradeoff between timeliness and accuracy of data. We all want data very rapidly and news quickly. But, obviously, that means you're going to have a smaller and less accurate sample, and the need for accuracy means that you're going to have to gather more information and process it more carefully, with greater quality control. That creates a tension.

A third element of data quality that is frequently addressed is cost. Publications have different budgets and have varying amounts of time, staff, and resources to devote to quality control.

Section 1. Analysis of the Database-to-Report Process

The Task Force analysis of the database-to-report process consisted of two phases: identifying steps in the database-to-report process and examining the database-to-report processes of four reports.

Identifying steps in the database-to-report process

The first activity of the Task Force was to examine the survey process as a whole to get a general idea of stages of the survey process from planning all the way through publication. In early meetings, the Task Force discussed some of the stages of this process and developed a flowchart showing the order and relationship of stages (Figure 1). This flowchart was used as a model to determine which specific steps would be examined in detail by the Task Force.

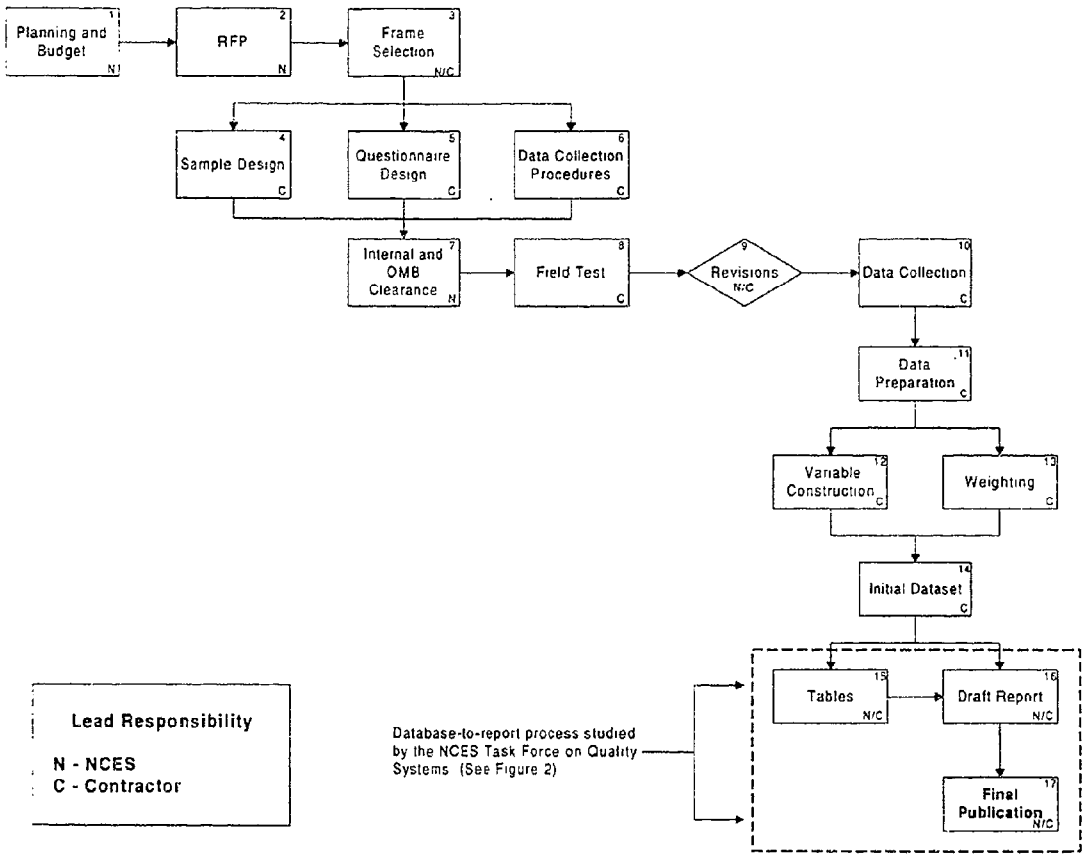
After reviewing the survey process as a whole, the Task Force decided to focus only on the final three stages: production of tables, writing of a draft report, and creating the final publication. There were three reasons why these steps were selected:

- a thorough review of the entire survey process would have represented an overly ambitious project with little chance of completion in a reasonable period of time;

- quality control procedures for early stages, such as sample and questionnaire design and data collection, already receive considerable attention and are generally regarded as thorough; and

- the errors discovered recently in NCES publications occurred late in the process, long after the datafile was completed.

Figure 1. Steps in the NCES Survey Process

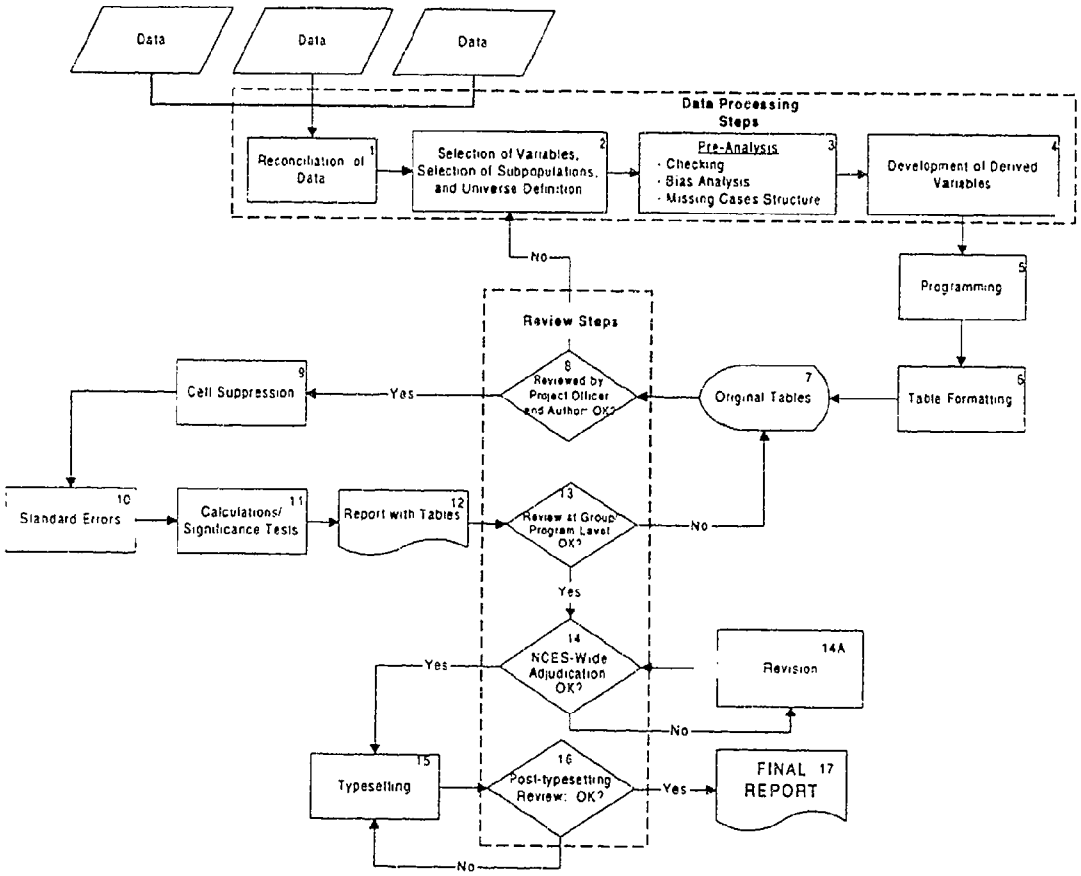


The Task Force determined that a review of these steps, which became known as the database-to-report process, would ultimately be most useful to senior NCES staff.

Turning to a detailed examination of the database-to-report process, the Task Force developed a second flowchart (Figure 2) showing the steps in greater detail. The flowchart was then reviewed and edited to give a general description of the final steps taken in producing NCES reports and tables. This model was used to develop questions for interviews and served as a starting point for studying the database-to-report process.

BEST COPY AVAILABLE

Figure 2. Steps in the Database-to-Report Process



NOTE: Though it is shown here as a single step, step 13 usually consists of two distinct parts: Program Review and Group review.

The database-to-report process can be broken down into 17 general steps, divided into three types of activities. Steps 1 through 4 are devoted to planning. Here, critical decisions (such as selection of variables) are made about how the data are organized, treated, and reported. Steps 5 through 7, 9 through 12, and 15 focus on production. In step 5 (programming) the decisions made in the planning steps are written into programs which produce the tables that serve as the basis for the report. In the remaining production steps, these tables are reformatted and edited; statistical analysis is performed; the text of the final report is drafted and integrated with the tables; and the entire report with tables is typeset and published. Finally, steps 8, 13, 14, and 16 are review steps, which traditionally have been expected to play the largest role in ensuring the quality of reports and tables.

All publications studied follow a basic course through the database-to-report process, which is made up of four major phases: preliminary table generation, table formatting and report production, internal review, and formal adjudication. The first major phase is the generation of preliminary tables, followed by a preliminary review of these tables. The first review step (step 8) is generally conducted by the contractor, the NCES Project Officer, or both. During this phase, the contractor and Project Officer review tables by checking against previously released reports and comparing current data with results from other sources. This review aims to determine the reasonableness of results and identify outliers and problem data cells.

The next major phase is production of the actual report. In some cases, production means largely formatting of tables. More often, text is written to accompany the tables. The draft report then undergoes a second, internal review (step 13), which is a combination of several reviews conducted either concurrently or sequentially depending on the needs of the program area. In this step in the sequential review process, the Project Officer reviews the draft report from both a substantive and statistical point of view. Once the concerns of the Project Officer have been addressed, the Program Director reviews the revised draft from both a substantive and statistical standpoint. Finally, when the Program Director is satisfied that all comments have been adequately addressed, the revised draft report is sent to the Group's Senior Technical Advisor for review. Once this review is completed and all revisions have been made, the draft is approved at the Group level and sent to the Office of the Commissioner to initiate the adjudication process of peer review.

The adjudication process is a uniform process across NCES programs intended to obtain internal and external peer review. Changes are made based on peer review, and the Chief Statistician or the adjudicator conducts a review before final approval to ensure that all changes have been made. Following the post-adjudication approval, some reports move directly into typesetting whereas some are edited by Information Services. If approval is not received, the report needs to undergo additional review at the Group level. (A more detailed discussion of review steps can be found in Section 3.)

While all NCES reports move through this basic process, there is a good deal of variation in terms of steps taken to move from one point to the next. The following sections analyze the database-to-report processes of each of five selected NCES publications.

BEST COPY AVAILABLE

The database-to-report process of four reports

1.1 National Assessment of Educational Progress 1994 Trends in Academic Progress

The majority of production work for this NAEP report is done by the Educational Testing Service (ETS), the NCES contractor for NAEP. The data processing and analysis steps of the database-to-report process are critical to the overall quality of the report. ETS begins the process by producing what it refers to as "almanacs" – a series of basic tables drawn from the datafile. ETS uses replication to ensure that the almanacs are error free, meaning that it runs multiple programs on the datafile and checks the results for consistency. These almanacs then serve as the basis for the remainder of production. Raw data are rarely revisited after the almanacs are completed. The almanacs serve as a basis for the cross-tabulations and significance tests that are used in preliminary tables. Once tables are produced, a draft of the report is written. This draft undergoes an internal quality review at ETS, in which all text and tables are checked against the almanac data. After changes are made as a result of this internal ETS review, the report is sent to NCES for Group review.

From this point, the report follows a standard path to publication. Changes made in the Group review are sent to ETS, where the report is revised and returned to NCES. Upon approval by the Program Director and the Associate Commissioner, the report enters the adjudication process. Following adjudication, ETS makes the final changes, checks the tables once more against the almanacs, and produces a memo outlining all changes made after adjudication. Once the report receives NCES approval through adjudication, ETS produces a camera-ready copy for the printer.

1.2 Digest of Education Statistics

The *Digest of Education Statistics* is compiled from a diverse array of data sources. Out of about 420 tables that appear in a typical annual release of the *Digest*, roughly 100 come from some print material, another 25 or 30 from unpublished sources, and the rest from raw data. These data come from all Groups of NCES, as well as 17 sources outside NCES. Because data come from several sources in a variety of formats, early steps in the database-to-report process of the *Digest* are necessarily quite different from those of other reports produced mainly from a single dataset. Staffing for the *Digest* is roughly equal to two full-time staff members at NCES and one for its contractor.

Responsibility for *Digest* tables is divided among two or three individuals. Each person must collect all information for datasets, set a timetable for production, and determine whether programming will be done in-house or with a contractor. Once these preliminary tables have been produced, they are checked for either programming error or possible misinterpretation of instructions first by the individual responsible for their production and then by the Project Officer. Any tables found with major errors, such as programming errors, are returned to the programmer, while those with minor errors are corrected by the reviewer. After this review, the Project Officer begins assembling the file containing all tables and notes. Following an optional review by NCES program directors to check for consistency between *Digest* tables and original data, the report enters into the normal NCES review and adjudication process, which for the *Digest* includes survey directors. Following adjudication, a final, detailed review is conducted by the Project Officer.

1.3 Schools and Staffing Survey in the U.S.: A Statistical Profile, 1993-94

For production of this report, the contractor, MPR Associates, played the primary role in the production of text and tables. The report was produced primarily by five people at MPR: three authors, one of whom also functioned as the overall coordinator; a research assistant responsible for some quality control, such as checking figures against previous reports and additional publications; and a programmer.

The process began with the development of a data analysis plan by MPR, which was then approved by NCES. After the approval, specific instructions were given to the programmer, and preliminary tables were generated. MPR first reviewed these tables internally, then NCES. The NCES review consisted of

- checking the tables against SASS estimates from previous years
- checking against estimates contained in *ED Tabs* from the same cycle of SASS, and
- comparing the SASS draft tables from MPR to draft tables received from other contractors for other projects.

Any changes requested by the Project Officer at NCES were made by MPR. After approval, preliminary tables went into formatting for inclusion in the final report.

At the same time, the original tables were used to draft the report, which was sent to the Project Officer for review. Here, results were again checked against *ED Tabs* and other available publications. After the report was approved by the Project Officer, it was integrated with the formatted tables and submitted for Program review, Group review, and adjudication. After all edits had been made and the report was approved, MPR produced the camera-ready copy, which was sent to GPO for printing.

1.4 Fall Staff in Postsecondary Institutions, 1993

For most IPEDS surveys, data are collected, entered, edited, and imputed by a single contractor, the Bureau of the Census. Once a final clean datafile is delivered, IPEDS staff are responsible for all aspects of the data-base-to-report process, including data production.

For this 1993 Fall Staff report, a relatively new survey, the Bureau of the Census collected, entered, and edited the data; then the edited file was delivered to another contractor, Westat Inc., for imputation, table generation, final report preparation, and production of the public use datafile.

In producing this report, NCES worked closely with its contractor Westat, which was responsible for generating preliminary tables and drafting the report. At both NCES and Westat, one individual oversaw data processing, while another had overall responsibility for the report. Once the preliminary tables and report were produced, they were reviewed first internally at Westat, then at NCES by the Project Officer, and later by the Program Director. Review steps at this stage included checks for logic and reasonableness, as well as editorial issues.

Following the review by the Program Director, the report was submitted for Group review and adjudication. Westat was responsible for making all changes and producing camera-ready copy, which was sent to GPO for printing.

1.5 Typical Database-to-Report Process for Common Core of Data (CCD) Reports

The database-to-report process for CCD is very much like other reports studied. Data are processed by a single contractor (Bureau of the Census). However, the contractor is responsible only for data processing and generating tables and not for any

further production of the report; instead, NCES staff are responsible for report production.

One interesting characteristic of CCD is the division of labor at NCES. Although a Project Officer is responsible for overall supervision of the report, two other supervisors are part of the team — one focusing specifically on the conceptual aspects and the other on the operational aspects. Once tables have been approved and formatted, the author assumes responsibility for writing the report, helping to move it through the standard NCES review process, proofing all text and tables throughout the review process.

Section 2. Typical Problems in the Database-to-Report Process

In examining the database-to-report process as a whole, three stages of the process appear to be the most error prone: production of preliminary tables, formatting of tables, and post-adjudication revisions. Errors are more likely to be discovered in earlier steps than later steps, as they have not been through as many review phases. In some ways, therefore, it is the later steps that present a greater overall threat to quality, as errors that may occur in late stages are less likely to be caught than those that occurred earlier.

There are four major types of errors in the production of reports and tables:

- Data compilation errors occur either as data are processed by NCES or the contractor, or as they are entered into databases before formatting.
- Inadequate documentation is the result of either a lack of information about how data were obtained or processed, or by different interpretations of labels or notation resulting in different utilization of data.
- Table formatting errors occur when data are put into the tables that are ultimately used in publications.
- Organizational characteristics affect the potential for error by either directly causing errors, or at least making them more difficult to detect and correct.

2.1 Data Compilation Errors

Data compilation errors are the most common of the four types of errors. These errors assume many forms, depending on the publication in question and the way in which NCES obtained data. For example, all of the data for the *SASS Statistical Profile* come from various components of the Schools and Staffing Survey, whereas the *Digest of Education Statistics* obtains data from some 17 sources outside of NCES, as well as almost every database within NCES.

Data that come in print format can be problematic because they must be keyed in manually, which is one of the most difficult parts of the process to regulate, as it is entirely dependent on the individual keyer. In addition to requiring detailed subject-matter familiarity, the quality and integrity of manually keyed data depend on the conscientiousness of the keyer. Even if data are keyed correctly, errors might have existed in the original publication; these errors are difficult to detect. Furthermore, in the case of certain datasets, it is difficult to locate mistakes if individuals are keying large amounts of data at one time.

Data that come in electronic spreadsheet format are subject to compilation or programming error. These spreadsheets often contain rows or columns of data calculated by some formula, but these formulas are not always manipulated or updated correctly. Verifying the data is difficult unless it is possible to cross-check against data from other sources; consequently, the likelihood of error increases. This type of error was recently discovered for a calculation of dropout rates that was to be included in the *Digest of Education Statistics*. In some cases, spreadsheet calculations are incorrect, but it is also possible that they are out of date, particularly with annual publications such as the *Digest*.

Recurring publications are particularly susceptible to some types of compilation errors in that some data structures change although most remain constant. Because no additional formatting or programming is required, many of the data items are updated from year to year using the same spreadsheet formats or programs, which generally reduces error and saves time. However, if the way that data are reported changes across years and this change goes unnoticed, data items can easily become lost or miscalculated because they do not fit with last year's format or program. In the production of the *Digest*, for example, data from outlying areas, such as for U.S. territories, regularly change from year to year as rows are removed or added. If data are inserted using the format from the previous year, certain elements will be lost because the structure of the spreadsheets will not match. In other cases, definitions of variables may change slightly from year to year, leading to errors in labels and footnotes as well as data inconsistency. Finally, such changes increase the risk that certain data items will not be revised as files are updated, or that older versions of the updated files will be used inadvertently instead of the most recently updated files.

Recurring studies generally have fewer errors than do one-time surveys and the errors that do occur are easier to identify. However, variable definitions may change from cycle to cycle, so new data may not reflect the same meaning as the old data. This change can cause confusion about how new and old data are to be reconciled and which structure is to be used. For example, a question asked in a 1972 survey may not have the same meaning as a similar question presented in the 1992 survey.

2.2 Inadequate Documentation

Documentation errors are the result of inadequate information about the data in a file or a table. There are two reasons why table information may be inadequate:

BEST COPY AVAILABLE

- there may be insufficient documentation of where the data come from or how they have been processed, and

- ambiguity about definitions of table labels can lead to inconsistent treatment of data in those categories by uninformed users.

Data in tables must be well documented. Information about how original tables were created is critical to the NCES review process. Several NCES staff members commented that it is difficult to pick up errors in such tables without some type of notation about how weights and variables were selected. Even if certain items in a table do not "look" right, it is difficult to pinpoint the problem without this information. NCES reviewers sometimes must go back to the programmer and review program code in order to determine how a table was generated.

Inadequate documentation includes data labels that do not clearly establish the context for a certain type of information, causing misinterpretation and error. Data may not be tabulated in a way that is consistent with how they were reported by survey respondents. Data users may be unfamiliar with NCES terminology and may misinterpret labels and titles that are not well documented.

Misinterpretation can also cause problems in working with contractors. Contractors and off-site programmers are often asked to generate tables but are not always given specific instructions about the exact weights, variables, and definitions to be used. As a result, these are sometimes chosen at the discretion of the programmer. Depending on the subject matter of the table and the level of expertise of the programmer, this discretion may be a cause of discrepancy if specific variables are not selected ahead of time. Sometimes this discrepancy is caused by programming error, but it also can result from miscommunication. Rarely does this lead to errors in final publication, as they are usually identified during review, but it does consume valuable review time which might be better used. Several individuals interviewed at NCES estimated that, for various reasons, roughly half the tables they receive from outside contractors need to be redone. Once again, this type of error is more common for publications where data come from multiple sources. For publications using a single contractor and a single dataset, it is easier to ensure consistency in the data labels and variables to be used.

Treatment of null data is an area particularly susceptible to documentation error. A zero or a blank space entered in a cell could mean that the item had no numerical

value, that the data were unavailable, or that data were not applicable for that cell. For every table or dataset collected from an outside source, null data may be interpreted differently, and interpretations are not always documented. Furthermore, statistical software applications such as SAS and SPSS can interpret blank spaces differently depending on default options, thus yielding different outputs depending on how null data are processed.

Recently, the *SASS Statistical Profile* and the *Digest* each produced tables showing private school principal average salaries, but there was a large difference (approximately $4,000) between the two reports. This difference resulted from different treatment of cells with zero salary in cells. The *Digest* included the cases with zeros, which in this case meant that the principal was not drawing a salary, whereas the *SASS Statistical Profile* excluded principals with zero salaries. In this case, neither assumption was necessarily erroneous, but the confusion about different figures could have been avoided if such assumptions were footnoted.

2.3 Table Formatting Errors

Errors in publications are often caused by changes in table formatting. Conceptually, these errors are not complex, but they are difficult to identify. Often, format changes that appear to be minor will have a major impact on data items. For example, in the case of one CCD publication the format required for the tables differed from the format of the data on the file. In attempting to reconcile the two formats, one column in the table was narrowed slightly, and in the process, the final zero was shaved off large observations in that column. The error was nearly impossible to detect because in the original data table the figure was correct. The shaving of the column occurred as the report was going to print, after the adjudication process. In a similar case, a directory of school districts was sent to print without phone numbers because the field which contained them was somehow eliminated during a final reformatting.

Another type of formatting error may occur when there are problems aligning data cells with labels. NAEP staff reported that formatting changes caused by review edits occasionally shift columns and not data labels, leading to errant observations. These types of error are common in early stages, but are generally corrected during the first quality check, because columns of errant data are more easily discovered than single observations.

2.4 Organizational Characteristics

Organizational characteristics also play a critical role in either increasing or reducing the potential for error. Three major organizational characteristics seem to contribute to error on different NCES publications, or at least make it more difficult to control it:

- channels of communication between publication staff and data sources,
- maintenance of "master" datafiles, and
- timing of publications.

Throughout NCES, there do not seem to be established channels of communication for Groups working on publications within NCES to communicate with each other, or in some cases, with contractors. Inconsistencies in publications sometimes arise because Groups within NCES do not have any way of knowing what other publications are currently in production within NCES, nor do they have the time to investigate such an issue. Consequently, teams or individuals working on a single dataset may treat the data differently and in turn publish different findings. Such a lack of communication was in part responsible for a recent error that nearly appeared in the *Digest*. In providing school count data to the *Digest*, CCD retained closed schools in the datafile for the first time last year. Due to a lack of communication about, and documentation of, this change, the *Digest* did not restrict its analysis to operating schools only, and, as a result, closed schools nearly appeared in the final count of schools.

Another potential contributor to error is the way in which master datafiles are created and maintained. On some publications, no single master datafile seems to contain all changes and updates to the data. Instead, several versions of datafiles are used simultaneously both within and outside of NCES, and there is often confusion about which datafile is most current. Furthermore, there does not seem to be a record of who actually accesses these datafiles or what changes are being made. This lack of a procedure for maintaining datafiles creates confusion about how to check and update them.

Finally, administrative pressures can contribute to error in two ways. First, as the importance of certain data increases or decreases, there is pressure to alter the format and emphasis of publications to reflect these trends. The more often data are moved and reformatted, the greater the risk for formatting error. This risk increases when restructuring occurs late in the publication process. Second, demand for certain

data occasionally changes or disrupts the time frame for the publication of reports. If particular information becomes immediately policy-relevant, there may be increased pressure to produce reports quickly. Responding to such increased pressure is one of the greatest dangers to quality control, because thorough review always takes time.

Section 3. Quality Control Practices at NCES

Overall, the Task Force found that there were no major systematic problems affecting data being released by NCES. There are quality control procedures in place at strategic points, and whereas some possible areas for improvement were identified, the general situation is very positive.

Given the vast number of opportunities for error to occur in the database-to-report process, it follows that the procedures for screening these errors ought to be very thorough. NCES and its contractors have developed extensive quality checks for all reports and tables. These procedures require a considerable investment of time and resources, and different Groups within NCES go about the process in a variety of ways. While technology plays an important role in developing the publication, the quality checks that are performed in the database-to-report process are almost always manual. Generally, such checks involve cross-checking tables against any of three types of sources: original tables, other reports that used the same data, or in the case of recurring reports, reports from previous years.

Major review phases

There are three critical review phases in the database-to-report process. Review phases are essentially a subset of the final two phases of the database-to-report process described in section 1.

- Phase 1: Early review (shown as steps 8 and 13 in Figure 2) is performed by contractors or by the NCES Group responsible for that particular publication.

- Phase 2: Group review and adjudication (step 14) is a standard procedure that all NCES data reports move through before publication.

- Phase 3: Post-adjudication review, occurs at the Group and in some cases at the contractor level.

The first and third review phases vary from report to report, while the middle phase remains uniform.

This section examines review procedures at each phase for all NCES publications. The text concludes with a discussion of some general characteristics of error-reducing practices.

3.1 Phase 1: Early Review Procedures

There are two critical steps in the early review process. The first step (step 8 in Figure 2) typically focuses on review of preliminary tables, while the second (step 13) reviews the completed draft of the report before it goes into NCES Group review prior to adjudication.

Procedures for reviewing and checking tables in the first step seem to vary in terms of the number of review steps and individuals involved. Preliminary tables can be checked in any of three ways: manual check of all tables against source data; add checks for rows and columns of tables, which allow for comparisons of totals and for check of consistency across tables; and key-and-verify, where the data are keyed twice and compared.

The second step of early review (a Program review of the draft report) is more standard, though there are some differences among reports. Decisions about specific review procedures at this point are made by the author, the Program Director, and the Senior Technical Advisor. The purpose of this step is to assure that the document is ready for peer review. This step is also the point at which it is determined whether the review will be "streamlined" or not. Streamlining is discussed later in this section.

For reports that rely mainly on a single data source, early review is conducted primarily by the contractor, while for compendium reports such as the *Digest*, a greater share of the burden falls on NCES staff assigned to the *Digest* and program area to review (usually under tight time constraints).

NAEP

As discussed in subsection 1.1, early review of the *NAEP 1994 Trends in Academic Progress* begins at ETS. The contractor prepares "almanacs" which are used to check the rough draft of text and tables. These quality checks are conducted manually by either a statistician or quality control support staff, depending on the complexity of the material and the amount of subject knowledge required. For this report, one individual at NCES is responsible for oversight throughout the review process.

Digest

As discussed earlier, tables for the *Digest of Education Statistics* are divided among several individuals. Each individual is responsible for generating a certain number of

tables, which are then assembled into a single file. Tables come from a number of sources. Each individual is responsible for the first review of all tables. This review is particularly important when tables are generated by contractors. It is estimated that about half of all tables produced by contractors need to be redone, generally due to programming errors or miscommunication between NCES staff and programmers. Depending on the type of errors discovered, tables are either edited and changed by NCES or sent back to the contractor. Once tables have cleared this first step in early review (the equivalent of step 8 in Figure 2), the Program Director conducts a cursory check to look for major errors. This phase of the review process is different for the *Digest* than for other publications because the publication contains little text. Consequently, once tables have cleared the first step in early review, they move directly to the Program Director for approval. Because data for the *Digest* come from so many sources, survey directors are also asked to check tables that pertain to their surveys for consistency, though this review is not mandatory at this stage of the process.

SASS

Early review for the report *SASS in the U.S.A. Statistical Profile, 1993-94* is conducted by the contractor (MPR Associates), who is responsible for generating preliminary tables. In addition, the tables are reviewed by the Project Officer at NCES. Once tables are produced, they are checked against previous SASS reports, draft tables from other reports and from concurrent publications, and other publications such as *ED Tabs*. The tables usually have gone through several iterations before they are sent to NCES for preliminary review, which is conducted by a Project Officer. Once approved, the tables go back to MPR for formatting. In the meantime, the report is written by MPR and reviewed (with preliminary tables) by the project officer at NCES, then merged with the formatted tables.

IPEDS

For the *Fall Staff in Postsecondary Institutions, 1993* report, preliminary tables are produced by Westat and pass first through an internal review similar to that conducted by MPR, where they are checked by multiple reviewers against other known data sources. Two persons at NCES are responsible for the report, one for data and the other for publication. Each has different review responsibilities. Preliminary tables are reviewed by the former, and are checked through replication. Once the preliminary tables have been approved, they are returned to Westat, where they are used to draft the report.

The draft report with tables then undergoes a technical and editorial review at Westat before being returned to NCES. This review includes overall logical checks to make sure that statements made about the data reflect sound understanding and methodology, and checks for correct use of terminology and footnotes. The draft report is then sent to NCES, where the Project Officer reviews the entire document.

Common Core of Data (CCD)

Early review procedures for this collection require involvement of multiple NCES staff. All data processing for CCD is contracted to the U.S. Bureau of the Census, which is responsible for generating preliminary tables. Once preliminary tables have been produced, they are sent to two internal reviewers who are in charge of conceptual and operational oversight of the report, as well as to an assistant who checks for technical soundness. Reviewers compare data with past years, across reporting levels for the same year, and within a single survey. They make value judgments about the plausibility of reported data. All comments are forwarded to the Program Director, who oversees the early phases of the database-to-report process. Edits are then sent back to the contractor, and this process is repeated until the Program Director has approved the preliminary tables.

The tables are then forwarded to the report author. The author is in charge of writing the actual report as well as shepherding it through the remainder of the review process. Once the report has been written, it is reviewed again by the Program Director and by the conceptual reviewer for approval.

3.2 Phase 2: Group Review and Adjudication

NCES Group review and adjudication is the second phase of the review process and is uniform for all NCES publications. Staff involved in all publications have identified this period as the most important for ensuring quality. During this review phase, decisions are made as to whether the draft publication is ready for review, and subsequently whether the document is ready for NCES publication.

Streamlined or full technical review

After the draft report has been produced, the author and Program Director, with assistance from the Senior Technical Advisor and the Associate Commissioner, discuss whether the draft will go through a streamlined technical review. In the streamlined

BEST COPY AVAILABLE

review, the Senior Technical Advisor reviews the report simultaneously with peer review rather than approving it prior to peer review. This decision is based on the complexity and sensitivity of the document, the newness and complexity of the database, and the experience of the author. The author, Program Director, and Associate Commissioner must all sign off on the decision to indicate the type of technical review to take place.

Group review

The next step in the process is the Group review, which determines whether the publication is ready for the peer review. In the regular technical review, the report must be approved ultimately by the Program Director and the Senior Technical Advisor. The draft may be reviewed simultaneously by the Program Director and the Senior Technical Advisor. The Associate Commissioner is also invited to comment on the publication during this phase of the review process.

If the streamlined review process is chosen, the Senior Technical Advisor's review can take place simultaneously with the peer review. The Program Director approves and signs off on the publication for simultaneous review by all subsequent reviewers. The advice of the Senior Technical Advisor and other staff may be solicited as necessary. In this review, the Program Director and author are responsible for ensuring that the data have been checked for accuracy and that the publication meets the NCES standards that guide technical work.

Peer review

Peer review provides a quality control measure by including reviewers from inside and outside the government who can contribute expertise, perspective, and diversity of opinions to the author. Peer reviewers include

- the Chief Statistician or designated adjudicator, who primarily is responsible for technical review and for ensuring adherence to NCES standards;
- a reviewer from the Data Development and Longitudinal Studies Group, who primarily is responsible for checking the data consistency in the draft against the NCES compendia and for raising any policy sensitivities that may arise;
- a reviewer from another program in NCES, who is primarily responsible for checking that the content of the publication is consistent with other research done by NCES, or if not, that differences have been fully explained; and

- one or more external reviewers, who provide additional subject matter or technical expertise and advice on the clarity of the publication.

In the streamlined technical review, the Senior Technical Advisor becomes a peer reviewer, since she or he has not had a previous opportunity to review the draft.

Adjudication Meeting

Once all comments have been received from the peer reviewers, the author recommends whether an adjudication meeting is necessary; however, the adjudicator makes the decision after all comments have been received. If the comments do not warrant major deliberations or revisions, the meeting may be unnecessary, although adjudication meetings rarely are deemed unnecessary.

The purpose of the adjudication meeting is to provide timely resolution of the peer reviewers' comments. The adjudicator chairs the meeting, during which the author presents the major points of the written comments. These comments are discussed by the participants, and an attempt is made to reach consensus on any differences of opinion. If no consensus is reached, the adjudicator makes the final decisions. Any appeals based on the decisions are made to the Chief Statistician, or subsequently, to the Commissioner.

After the adjudication meeting, the author makes any necessary revisions to the document and resubmits the draft with a post-adjudication memo to the adjudicator, which lists how comments from the reviewers have been addressed. The author and adjudicator then jointly decide whether all necessary changes have been made to the publication. The Chief Statistician next approves or disapproves the draft, based on the adjudicator's recommendations.

3.3 Post-Adjudication Review

Once changes have been made following the adjudication meeting, the report is returned to the adjudicator for final post-adjudication review and sign-off. Then, it is forwarded to the Chief Statistician for final approval. After this final approval, the report is considered ready to publish from a procedural standpoint.

A final check for the *SASS in the U.S.A. Statistical Profile, 1993-94* has recently been developed by MPR. Following Chief Statistician approval, any changes made throughout the review process that have required programming are used to update the

programming instructions. The programs are then run again based on the updated instructions, and the outputs cross-checked with the final report. This additional step has helped to find errors that had previously gone undetected.

Post-adjudication review procedures for the *Digest* are even more comprehensive. Following Chief Statistician approval, the Project Officer performs a detailed review of several tables. Tables that are known to be error-prone are cross-checked, line by line, with those in the original database. Other random checks are also performed in the same manner. This review is intended to be a final, catch-all process. Additional errors are often discovered at this point.

3.4 Additional Quality Control Measures

Reviews are an important part of quality control for NCES publications. In addition to review, however, NCES Groups and contractors use a number of strategies and techniques to further reduce error. Some of these strategies are specific and technical, while others reflect a general approach to quality control procedures that cannot be described in procedural terms.

In spite of all of the extensive measures employed to reduce error in NCES publications, there is still no substitute for subject knowledge and experience with a particular publication. When reviewing large quantities of data, it is impossible to perform detailed reviews of every single cell in every table. The most important skill in reviewing reports seems to be knowing where to look and what to look for in terms of possible errors. These skills generally only come with experience, but in some cases, steps have been taken to share that experience at a Group-wide level. For example, the Project Officer of the *Digest* recently assembled a list of specific checks for tables that tend to be susceptible to error. This technique allows the experience of one individual to be used by less experienced staff and gives senior staff more time to review particularly complicated sections of reports and tables.

Another common practice for preventing error is the documentation of all changes made to reports at each review stage. Different Groups and contractors do this in different ways. As discussed earlier, all authors produce a memo outlining all changes made after adjudication. MPR, in addition, maintains a card file describing all changes made throughout the process.

Finally, automation at certain points of the process seems to play some role in minimizing error, though that role appears to be limited. Both MPR (SASS) and Westat (IPEDS) transfer most of their tables into typesetting electronically. This is done by creating standard formats and shells for tables that can be reused from year to year. Electronic processing of tables minimizes the amount of formatting and manual editing required, thus helping to reduce error.

Section 4. Production of Reports at Other Federal Statistical Agencies

To better evaluate the quality of NCES publications, the Task Force contacted other Federal government statistical agencies. Learning about quality control at other agencies presented a challenge since each agency has its own system of quality control and seldom is this system centralized and well structured. However, the reports below do give some perspective within which the quality of NCES publications can be viewed.

4.1 Bureau of the Census: *Statistical Abstract*

The database-to-report process of the *Statistical Abstract* most closely resembles the *Digest* in terms of its structure and procedures. Like the *Digest*, tables for the *Abstract* are divided among a few statisticians who are responsible for their production. Data for the *Abstract* come from over 250 sources.

The process begins when the Bureau of the Census sends out a letter and form requesting updated information from all sources. Data come back in several forms. In some cases, the completed form sent out by the Bureau of the Census is returned and data are filled in accordingly. Sometimes, they receive a copy of a publication and must key the data manually. In other instances, the data come in electronic format. Finally, some data are simply downloaded off the Internet.

All print data are first keyed manually into a worksheet, then verified by a second person. All electronic data are processed by one of the four statisticians, who are jointly responsible for the creation of all tables.

Unlike NCES, there is no step in the database-to-report process at the Bureau of the Census that equates to the adjudication process. All review is handled internally. Review consists of manually checking tables against original data sources, comparison to tables from previous years, and inserting "add checks" (extra rows and columns which provide summary figures) into worksheets to check totals. Once all tables have been reviewed, support staff insert typesetting codes and convert the worksheet files into ASCII format. The *Statistical Abstract* is produced using XYVision software, which automates all formatting steps. Once typesetting codes have been inserted, the human factor is essentially removed from the process. Proof pages are produced, which are then reviewed by statisticians and editorial assistants, and checked against original tables or data sources.

While there is no specific documentation of changes made to the *Statistical Abstract*, the database is set up in a central location in such a way that it cannot be edited by more than one individual at any time. As a result, a record is maintained of the last date that a staff member updated the datafile.

Review procedures for the *Statistical Abstract* closely resemble those of the *Digest*. There are, however, some significant differences. First, all data that are keyed manually into tables are cross-checked with the original print data. This is not a technical review, but rather a preliminary search for manual error. Both for the *Digest* and *Statistical Abstract*, survey directors have the option of reviewing tables and proof pages, though this is not required.

In the last 7 years, the Bureau of Census has taken extensive steps to automate the production of the *Statistical Abstract*. Once tables are reviewed, they are coded for electronic typesetting. The coding system is very sensitive to error, so if a code is entered incorrectly, it is obvious when the report is typeset. This automation has helped greatly in reducing format errors, particularly in the final stages of production, because the human element is removed from the process.

4.2 Bureau of the Census: Demographic Statistical Methods Division

There does not exist a Bureau-wide office of quality systems responsible for setting standards and procedures. Rather, areas within the Bureau develop quality control procedures that are customized to that area's needs. One general approach that has worked is to computerize as many processes as possible eliminating the human intervention that is expensive, time consuming, and a source of error. Several processes were reviewed with this idea in mind. For example, humans had been used to transcribe data from one computer system to another. Computerizing this operation significantly reduced errors.

The review procedure for reports is similar to that used at NCES whereby analysts, that is, authors, check with statisticians before the report can be published. Although formal channels and procedures exist for signing-off, a formal adjudication process is not in place, and no formal Bureau-wide strategies exist to check for and eliminate both data and formatting errors in the document.

To further reduce error, training programs have been implemented. For example, a manual has been prepared and regular courses offered to help staff write to specifications.

With respect to multiple checking, this is not considered to be a reliable method of reducing errors. This follows the teachings of Prof. E. Deming, who believed that multiple checking of documents was basically a waste of resources and one of the least likely strategies for improving data quality. The argument is that increased checking leads to a situation in which nobody feels responsible for the product.

4.3 National Center for Health Statistics

Quality control procedures in NCHS are developed and applied independently and separately within each of the data collection programs. Because NCHS operates a wide range of data collection systems and produces a large number of publications and because many of these programs are relatively small, the number and variety of quality control procedures are large. NCHS does have, however, a systematic review process. All publications are reviewed through the supervisory chain within the data collection program. A final review is conducted within the office of the Center Director but the vast majority of quality control activities are undertaken within the programs themselves. All publications also undergo editorial review in a central office.

Special attention is paid to results that appear in several publications and considerable efforts are expended to ensure consistency. Errors in publications and errata sheets are not rare, although the rate is low relative to the total output of data and publications. It is usually easier to locate errors in recurring publications (e.g., vital statistics), while one-time publications and publications require more effort.

4.4 Bureau of Labor Statistics: Publications

Although BLS does have a unit called Quality Management Staff, its role is to assist in quality improvement efforts rather than setting and enforcing Bureau-wide policy. Each program within the Bureau tends to develop its own procedures for maximizing quality. Staff are encouraged to resort to machine edits, register period-to-period time changes, and to research changes that exceed a certain threshold. Fortunately for the Bureau, most of the surveys are recurring, which facilitates error-identification and correction.

Once data are final, manual, review-based quality control efforts begin. Manual review by experts is a popular strategy. The focus is on areas with errors in the past and on estimates that are particularly sensitive or that have a disproportionate effect on national and state averages. Clearly, there is a potential for human error and little is known about how to efficiently reduce, much less eliminate, this source. Errors, however, are rare although there have been problems with external pressure to release data prematurely.

In the Publications Office, the challenges are very similar to the ones on which this Task Force has focused. The main strategy is to compare current results with previous results in the same series of publications. Also, special attention is placed on outlier units, such as commodities, that account for a large portion of the national or regional total. Procedures tend to be relatively simple with few common standards, although there is an intention to begin developing standards. Problems do occur when several versions of the same program, data, or tables are used concurrently. The challenge is to institute a system that keeps track of the changes and ensures that only one version is used, usually the latest one. In some cases, the error-detection mechanisms are built into the system, primarily because so many of the publications are repeated, sometimes on a weekly basis.

4.5 Bureau of Labor Statistics: *Current Price Index*

An explicit quality program at BLS was designed to maximize the quality of the final output of the *CPI*. Errors were reduced to virtually zero through a fairly simple but effective combination of thorough and repeated checking, comparison with external sources, and careful control of the data processing system.

In addition, a study of the review process indicated that reviewers at the early stages were passing on responsibility to reviewers at later stages. This shift was in part due to the early reviewers' realization that there was much overlap between the stages. A thorough evaluation of this process led to very specific delineation of objectives and responsibilities at each stage.

4.6 National Agricultural Statistics Service

NASS has a Technical Review Program which involves not only the head office in Washington but also the 45 field offices. Since many of the reports are repeated,

BEST COPY AVAILABLE

much of the review strategy is based on a comparison across time. Dramatic changes are flagged by the computer system for further investigation.

Results are reviewed by the Agriculture Statistics Board which applies "expert judgment" before the data are released. These experts will question the soundness of the results and will call for further investigation if needed.

NASS averages more than one publication per day. Time pressure sometimes results in incomplete reviews, and errors have been known to creep into the results. More efficient ways of processing data are being researched to allow more time for the review process.

4.7 Statistics Canada

Statistics Canada does have quality control standards and procedures but they are not necessarily formal, institutionalized, and applied uniformly across the agency. Although there is a group of professionals with responsibility for quality control issues, their effort is more reactive than proactive.

Each area in Statistics Canada has autonomy for performing quality control, and each decides independently how to validate the data. Generally, common sense procedures are applied with focus on big changes, big contributors, historical data, and related variables. Autocorrelated processes are also studied and process control charts are sometimes used in the search for errors.

One interesting exception is in the case of clerical staff which recently was brought together under one administrative umbrella. (Previously many of the areas within Statistics Canada had their own clerical operations and staff.) A quality control system was developed for this new entity and is being successfully maintained.

Also, some problems recently identified in CANSIM, an integrated database containing a wide variety of national and regional statistics, will likely result in a review of procedures used in preparing data for this database.

II. RECOMMENDATIONS BASED ON FINDINGS

The Task Force found that the quality control procedures utilized at NCES vary somewhat throughout the agency. While standardization of procedures is complicated and may present problems, we would suggest that at least some of the following recommendations be applied uniformly throughout NCES.

As is argued elsewhere in this report, quality is a multidimensional concept made up of timeliness, cost, and accuracy. In some cases, changes in operations may improve one aspect of quality without necessarily detracting from another. In other cases, however, efforts to decrease error might result in a decrease in timeliness or an increase in cost. When implementing the Task Force's recommendations, NCES should decide how much more time and resources it is willing to devote to ensure the quality of its data.

Finally, in considering the following recommendations, one should distinguish between "preventive" and "curative" procedures. Recommendations addressing error prevention should be given higher priority since preventive measures typically occur earlier in the database-to-report process and are usually, although not always, less expensive. The importance of preventing problems cannot be overestimated, in that doing so can considerably simplify subsequent steps in the survey process, including quality control. Serious investment in preventive quality control measures can often be recouped through simplification of the later stages of data processing and review.

Section 1. Themes of the Recommendations

The following list of "recommendations," rather than representing a definitive set of prescriptive recipes, resembles more an extensive list of ideas that the Task Force considers to have potential benefit. These ideas, categorized into five general themes, are a first step in the process of looking at a variety of options that might further increase the quality of NCES publications. From this list, the Task Force selected six final recommendations as being most viable. Those final recommendations are highlighted on the following list and discussed later in this report. Recommendations followed by an asterisk might better be implemented at the Program level rather than throughout NCES, mainly because of operational differences across programs.

Spread Existing Skills Within NCES

- *Develop a handbook of technical guidelines for NCES publications*
- Train appropriate NCES staff in statistical software applications

Improve Documentation and Communication

- Front end
  - *Include input from survey staff*
  - *Improve communication between NCES staff and programmers*
  - Improve documentation of data used to create tables
  - Improve documentation of database status
  - Improve consistency across publications*
  - Improve review criteria for checking tables*
- Back end
  - *Formalize internal documentation of tables**
  - Improve archiving of NCES survey datafiles

Improve Uses of Computer Systems for Checking Key Results

- Generate supplementary cross tabulations*
- Develop master analysis files for a report*
- Provide additional programming support for the Chief Statistician

Improve the Review Process

- *Increase use of contractor help in the review process*
- *Increase NCES staff available for review*
- Define objectives and responsibilities for each review step
- Delineate clearly activities, responsibilities, and timing for each review step
- Share specialized review procedures throughout NCES

Improve Contract Procedures

- Allocate specific funds for data quality in contracts*
- Improve the use of data analysis plans*

Section 2. Findings and Preliminary Recommendations

This section presents each of the themes listed previously, the findings of the Task Force, and a discussion of the preliminary recommendations for that theme.

Spread Existing Skills within NCES

Findings

- There is a great deal of variation across programs with respect to the formats, definitions, labeling, and technical guidelines used to produce of reports and tables.

- Senior staff time is not being used optimally in the review process, as senior staff spend too much time on routine quality control activities, such as checking the consistency of data cells across tables, that could be undertaken by more junior professionals.

Recommendations

1. *Develop a handbook of technical guidelines for NCES publications.* The handbook, intended to supplement the existing *OERI Publications Guide*, would contain suggested formats, labels, and definitions for NCES publications. Such a handbook would help NCES to develop more consistently documented tables with standardized labels. The handbook could also contain examples of good titles, table layouts, and descriptors. (The handbook would be far more detailed and NCES-specific than the *OERI Publications Guide*.)

2. *Train appropriate NCES staff in statistical software applications.* Training should be available for all staff, and mandatory for some, in the use of statistical software packages, such as SAS and SPSS, and spreadsheet software, such as Lotus and Excel. The objective would be to establish minimal cross-tabulation and table design skills at additional points in NCES so that Project Officers and authors could do their own computer runs on their survey data without having high level computing abilities. Such expertise would allow early identification of any errors being produced by contractors. Generic programs specific to surveys could be developed, whenever feasible, so that survey staff would have to change only variable names to generate cross-tabulations. Frequent changes in programming-

support staff make checking important. Independent replication of selected tables could thus be used to verify programming code.

Improve Documentation and Communication

Findings

- When secondary analysis is being carried out, survey staff are not always consulted.
- There is room for improvement in the documentation of tables regarding various aspects including the selection of weights and variables. Inadequate documentation makes review of tables more difficult within NCES and, even more so, by outside users.
- In addition, there is some evidence of inconsistent reporting of the same result in different publications.
- There is some lack of consistency in how certain commonly occurring situations are treated in tables, such as handling zeroes and whether users are warned about less reliable data.
- Poor communication with programmers can lead to misunderstandings that consume a considerable amount of time and resources in the generation of tables. For example, programmers often make substantive decisions, based on inappropriate assumptions, which are not always fully discussed with NCES staff.
- Distance and infrequent contact between programmers and NCES staff may contribute to these problems.
- Internal communication procedures regarding modifications to NCES databases need to be improved. In particular, it is sometimes not clear which version is the final version, where it is located, and what changes have been made.
- Some survey datafiles are subject to updating over an extended period of time. This can create problems when master files are updated after reports have been prepared using earlier data since now the file used to create the analysis no longer exists.
- In addition, some historical databases are currently archived only on a mainframe. This is not an adequate archiving procedure. These archived files also frequently suffer from inadequate documentation.

Recommendations

3. *Include input from survey staff when doing secondary analysis.*
   In cases where secondary analysis of data is being performed, survey staff knowledgeable about the data could be included in the publication development process.

4. *Improve communication between NCES staff and programmers.*
   Recognizing that this relationship varies across the agency, some general suggestions for improving communications include:

   - Encourage NCES staff to prepare clearer and more complete specifications for tables and data runs, including information about items such as variables, weights, and versions of the datafile that should be used.
   - Encourage programming staff to ask more questions when specifications are not clear or complete.
   - Use various telecommunications media such as telephone, fax, E-mail, and Internet to communicate more frequently and efficiently with off-site programmers.
   - Arrange for contractors to have representatives readily available on a regular basis by visiting NCES frequently, by working on-site, or by being easily reachable through other means such as telephone, fax, and E-mail.
   - Arrange for NCES staff to visit off-site programming sites.

   Better communication between programmers and NCES staff should help reduce the need for programmers to make unilateral substantive decisions without fully consulting with the Project Officers and authors of publications.

5. *Improve documentation of data used to create tables.*
   As part of its backup documentation, every table should have the exact name and version number of the datafiles used to create it.

6. *Improve documentation of database status.*
   It is imperative that a record be maintained of how databases have been updated over time. NCES should develop a system for documenting

   - where the latest version of a database is located,
   - who was responsible for any recent modifications to the database,

BEST COPY AVAILABLE

- when the modifications took place, and
- what type of modifications were made.

In addition, standardized procedures should be developed for the updating, sharing, storing, and ultimately archiving all databases.

7. *Improve consistency across publications.*
   Procedures should be developed that identify in the early stages of a publication any other NCES publications that might deal with similar concepts, and investigate and resolve if possible, or at least document, any inconsistencies between those publications prior to submission for adjudication.

8. *Improve review criteria for checking tables.*
   Criteria for reviewing tables need to be specified in advance and shared with the contractor or whoever is responsible for creating each table. Specifying the criteria in advance would reduce the number of edit cycles needed for these reviews.

9. *Formalize internal documentation of tables.*
   Full documentation of the weights and variables selected should accompany tables being prepared for NCES publications. It also would be valuable to append or link complete versions of original programs, together with documentation, to the files containing tables, so that the programs would be part of the historical files for that publication.

10. *Improve archiving of NCES survey datafiles.*
    Historical datafiles currently archived on the mainframe or on magnetic tape need to be stored on CD-ROM or otherwise protected to avoid deterioration. These files also could be archived with the Inter-university Consortium for Political Science Research (ICPSR) for long-term storage and for downloading when needed.

Improve Uses of Computer Systems for Checking Key Results

Findings

- Certain types of review may be improved (in terms of error reduction, timeliness, and use of staff resources) by making better use of computer resources to assist in the simple checking of tables.

- Procedures and systems for checking tables for NCES publications vary considerably within NCES. However, some practices that have proven to be particularly effective for certain publications were identified.

Recommendations

11. *Generate supplementary cross-tabulations.*
    Each survey program area might produce key data tables as benchmarks for other data users. Additionally, more CD-ROMs could be generated, similar to NAEP almanacs with voluminous cross tabulations already calculated, though it is imperative that such tables must be thoroughly checked if they are to be used as benchmarks. Staff could use the tables to check totals on various publications that could eliminate the need for new additional tabulations. Using these tables would also reduce the need for National Education Data Resource Center (NEDRC) or NCES staff to produce special tabulations in response to information requests and would also provide a product for the NCES Web site.

12. *Develop master analysis files for a report.*
    Some publications have a strategy whereby data are first prepared and finalized in a form that is internally consistent and incorruptible. For example, the *Digest* sets up all data in a spreadsheet whose cells are linked to maintain consistency of all tables utilizing the same data. Incorrect entries are flagged when they cause inequalities between cells. Such a strategy might be fruitfully adopted by other publication teams.

13. *Provide additional programming support for the Chief Statistician.* The Office of the Chief Statistician should obtain computer programming support to spot check the programming for incoming publications.

Improve the Review Process

Findings

- NCES has in place a review process that is thorough until a report is approved by the adjudicator. The Task Force has identified four separate review activities that are undertaken within NCES: Project Officer and author review (step 8 of Figure 2), the Group/Program review (step 13), the adjudication process (step 14), and the post-typesetting review (step 16).
- Not all review activities are completely standardized across groups.
- Considerable overlap exists in the reviews done by various staff members and at various stages of review.
- Currently, insufficient NCES staff are available to carry out additional reviewing of publications; furthermore, the situation has been deteriorating and is not likely to improve soon.

Recommendations

14. *Increase use of contractor help in the review process.*
As an alternative or in addition to increasing NCES staff, NCES might use contracted staff, such as ESSI, to review some types of publications or to provide a fresh look at some publications.

15. *Increase NCES staff available for reviews.*
Additional resources should be added to the Group technical reviews. Currently, three staff members are assigned the bulk of the responsibility for reviewing publications at this level. On the whole, fewer than three full-time equivalents (FTEs) are assigned to this review. Allowing junior staff to handle routine quality control procedures could reduce the burden on senior staff and increase the timeliness of reviews, particularly for routine publications.

16. *Define objectives for each review step.*
Currently, all aspects of a publication are usually reviewed at each of the first three review steps. NCES management might consider designating specific goals for each review step in order to maintain the overall accuracy of publications while reducing the amount of time spent in each review. Doing so could potentially reduce the time and cost involved in producing publications, thereby improving their quality.

This is an NCES-wide recommendation that should be addressed by management. The goal would be to reduce overlap in objectives between the various review steps.

17. *Delineate clearly activities, responsibilities, and timing for each review step.* Once goals have been set for each review step, NCES should specify the type of activities that need to take place at each of the four steps. The goal would be to significantly reduce the overlap in review responsibilities that currently exists in the review process. Such specification could result in authors and managers taking more responsibility for assuring the quality of their publications and data products, rather than either relying on the formal review process; or in having the steps be more simultaneous than sequential.

18. *Share specialized review procedures throughout NCES.* In addition to clarifying the review process, some NCES procedures now used for some reports could improve the quality (more timely, fewer errors, and less cost) of NCES publications if they were used more consistently throughout NCES. These procedures should be identified, assigned priorities, and developed.

Improve Contract Procedures

Findings

- Certain quality control practices currently used by NCES and its contractors might be better implemented if such procedures were stipulated in contracts.

Recommendations

19. *Allocate specific funds for data quality in contracts.*
An explicitly defined portion of funds should be set aside within each analysis task budget for data quality activities. Contractors should be required to

    1. submit a data quality plan with any proposal, focusing on specific quality measures to be implemented by the contractor;

    2. implement the plan throughout the production of the report, and

    3. show how quality control measures were implemented when the report is submitted to NCES for review.

20. *Improve the use of data analysis plans.*
Although data analysis plans are routinely prepared at NCES, there may be a need to review how these plans are prepared, their level of detail, and the emphasis that staff place on them. In particular, greater attention to variable definition and specification during the planning stage might help to reduce the errors in tables and the time currently spent on checking them.

Section 3. Final Recommendations

None of the recommendations will have an impact unless there is agency-wide support for them and for improvement of quality. Not only Program Directors but authors as well should take responsibility for final products, especially reports. Errors should be corrected as close to the source as possible with the author (or Project Officer if the author is external) taking as much responsibility as possible for the quality of the final prod .

The Task Force evaluated the 21 recommendations, along with an additional recommendation proposed during the January 1997 NCES meeting, to select those that were most feasible and potentially beneficial. As a result of this process, the Task Force eliminated 2 of the 22 recommendations and selected six of the remaining 20 for further development. The arguments for selecting these 6 are presented below. (Two of the recommendations are sufficiently similar to warrant joint discussion.) Much of the text that follows is taken directly from the previous section.

1. Develop a handbook of technical guidelines for NCES publications

The handbook, intended to supplement existing *OERI Publications Guide*, would contain suggested formats, labels, footnotes, and definitions for NCES publications. Such a handbook would help to standardize able style and help NCES to develop more consistently documented tables. The handbook could also contain examples of good titles, table layouts, and descriptors. (The handbook would be far more detailed and NCES-specific than the *OERI Publications Guide*.)

The handbook would also include suggestions about organizational and administrative procedures that might result from subsequent quality control activities. For example, the handbook could include the recommendation that publication staff look for and include input from survey staff. This would be one of a list of generally approved approaches that have been shown to be beneficial in certain situations.

Keeping the handbook in a three-ring binder would be used to allow it to be a dynamic tool that would change as practices and procedures improved with time and experience.

BEST COPY AVAILABLE

The Task Force believes that this recommendation will reduce the number of errors and the time spent reviewing individual tables. If reviewers can worry less about the superficial aspects of the tables, they can devote more effort to reviewing the content.

The Task Force recognizes that this recommendation will not be implemented quickly and that additional resources will be required to put the process into place and to maintain the necessary momentum. The suggestion is that a group be assigned to develop the first version of this handbook. The group would include the current Senior Technical Advisors (Mary Frase, Andrew Kolstad, and Marilyn McMillen), who would require support staff to help put the document together. In particular, participation by staff who prepare reports for review would be encouraged. After the initial version is developed, long-term responsibility for maintaining the handbook would be assigned to one person who would periodically arrange for its updating.

2. Include input from survey staff

In cases where secondary analysis of data is being performed, survey staff knowledgeable about the data should be included in the publication development process. This recommendation refers only to internal NCES staff.

The Task Force saw many examples in which the input from survey staff would have helped avoid problems and quickly resolve any apparently inconsistent data.

Furthermore, the Task Force believes that this recommendation is not very expensive or difficult to implement. First, it should be made widely known who the contact person is for each survey. Second, the publication staff should be encouraged to include the survey staff contact person in all phases of the publication process, especially during the design stages. This would be done by talking to staff and Program Managers and by including this item in a checklist in the handbook (see previous recommendation). Third, it needs to be emphasized to survey staff that it is part of their job to cooperate with analysts doing secondary data analysis.

3. Improve communication between NCES staff and programmers

The Task Force believes that this recommendation also has potential benefit at relatively little cost, although it does take effort at the outset in writing detailed specifications.

Recognizing that the staff–programmer relationship varies across the agency, here are some general suggestions for improving communications:

- encourage NCES staff to prepare clearer and more complete specifications for tables and data runs, including information about items such as variables, weights, and versions of the datafile that should be used.
- Encourage programming staff to ask more questions when specifications are not clear or complete.
- Use various telecommunications media — such as telephone, fax, E-mail, and Internet — to communicate with off-site programmers.
- Arrange for contractors to have representatives readily available on a regular basis by visiting NCES frequently, by working on-site, or by being easily reachable through other means such as telephone, fax, and E-mail.
- Arrange for NCES staff to visit off-site programming sites.

Better communication between programmers and NCES staff should help reduce the need for programmers to make unilateral substantive decisions without fully consulting with the Project Officers and authors of publications.

Lines of communication need to be improved in both directions. On the one hand, NCES staff need to make every effort to provide programmers with clear specifications, and, on the other hand, programmers need to question NCES staff when specifications are problematic in any way.

The Task Force suggests that a group be set up to study this topic and decide what steps should be taken to implement the recommendation.

4. Formalize internal documentation of tables

Full documentation of the weights and variables selected should accompany tables being prepared for NCES publications. It also would be valuable to append or link the complete versions of original programs, together with documentation, to the files containing tables, so that the programs would be part of the historical files for that publication. The Task Force believes that this recommendation will significantly shorten the time needed to review individual tables. (It will, however, lengthen the time needed to prepare them.)

At one point, the Task Force considered including a more complete set of notes to cover items such as variable definitions, selection criteria for subpopulations, and recoding rules. However, the Task Force decided that this would be too onerous a task and should not be part of the recommendation especially since this information could be obtained from the attached or linked programs, albeit in a less accessible form.

The implementation of this recommendation would involve a small group meeting a few times and setting up procedures and criteria to decide what should and should not be attached. As program needs are likely to vary, specific attention should be given to the procedures and criteria of each publication.

5. Increase NCES staff available for review and increase use of contractor help in the review process

Additional resources should be added to the Group technical reviews. Currently, three staff members are assigned the bulk of the responsibility for reviewing publications at this level. Although a new contract support person has been added in one group to assist in this work, all members have other major responsibilities. On the whole, fewer than three FTEs are assigned to these reviews. Allowing junior staff to handle routine quality control procedures, would reduce the burden on senior staff and increase the timeliness of reviews, particularly for routine publications. These junior staff members would have responsibility for activities such as checking numbers, reviewing grammar and style, and ensuring consistency within the report. Contractors could also be used to help in activities such as word processing and editing.

As an alternative, or in addition, to increasing NCES staff, NCES might use contracted staff, such as ESSI, to review some types of publications or to provide a fresh look at some publications.

The Task Force believes that this recommendation warrants special attention on the part of the Commissioner because it involves assigning additional resources to this task. However, it is argued that increased diffusion of responsibility for reviewing would allow more timely reviews and release the bottleneck that sometimes occurs at this stage of the publication process. Earlier and more thorough reviews would result in less deliberation at later stages and less rewriting.

Management Participation

The success of any quality improvement effort requires strong, active, and continuous support from top management. This means "walking the walk," not just "talking the talk."

The key is for the top management (such as Commissioner and Associate Commissioners at NCES) to be personally involved. Some of these top managers might participate as members of a quality improvement team (QIT), whose objective might be to determine, for example, how best to make more staff (both contractor and NCES) available during the review process. Top managers not on the QIT should discuss the progress of the teams with their staff who are team members. These staff members should be encouraged to report to the other members of their Group on the team's progress and to ask for input from other interested staff members.

Every six months, the top managers should review the list of quality improvement recommendations made by the Task Force, supplemented by recommendations that are made by the QIT. For which of the recommendations have improved procedures been identified? Have these procedures been implemented? If implemented, have they been used? If used, have they been found to be successful? Finally, are resources available to allow progress to begin on the other recommendations that have not yet been addressed?

We would encourage management to be personally involved in quality improvement efforts, to ask questions of staff that demonstrate management concern, and to periodically review progress to date. In this manner, management will demonstrate its concern that will motivate NCES staff members to recognize that everyone is working to maintain NCES' reputation for being a source of information that is of the highest possible quality. Quality is an issue of culture in the organization, and it is the responsibility of management to ensure that this attitude is pervasive at all staff levels.

The Task Force recommends that a general QIT be created in order to ensure that the recommendations of this report are followed up and that the issue of quality remains a permanent topic of concern at NCES. The team need not be large, but it should include members from several sections of the agency, including senior staff. The team needs to meet fairly frequently and, while it should not become involved in extensive tasks, it should have enough visibility and presence to oversee the

implementation and maintenance of sound quality control practices in NCES. It is important that resources be allocated to this activity.

Section 4. Future Work

The earlier sections present and discuss a set of recommendations whose implementation should be considered further by NCES Senior Staff and Program Directors. Furthermore, the Task Force has identified a few other areas that may benefit from further study.

As is stated in the introduction, the Task Force decided very early in the process to limit the review to the database-to-publication process. Now that this process has been thoroughly reviewed, NCES might want to discuss the usefulness of considering other, earlier, stages of the survey process such as creating the database, questionnaire design, data collection procedures, etc.

There is also the possibility that additional lessons could be learned by a more thorough review of practices that have been implemented in other organizations. This would include not only other government statistical agencies but also research groups in the private sector and in academia.

This report focuses on hard copy publications. However, NCES also disseminates data in a variety of other formats: CD-ROM, diskettes, Internet, and ad hoc requests. It might be important to study quality systems used in the preparation of these products.

Finally, some attention might also be paid to information being released in more informal and direct means, such as the NEDRC and the Web site.

In summary, areas the Task Force suggests warrant further investigation are

- earlier stages of the survey process,
- general quality control procedures in other organizations,
- quality of information released on CD-ROMs and other electronic formats,
- NEDRC and the quality of the information released by this office, and
- the NCES Web site and the quality of the information on this site.

APPENDICES

Appendix A. Memo on New Quality Control Procedures

TO: Jeanne Griffith
Associate Commissioner for
Data Development

FROM: Tom Snyder
Program Director
Annual Reports

SUBJECT: New Quality Control Procedures

Conceptual Thoughts on Quality Control

Although we are careful in our production of statistical reports, some error seems inevitable. By definition, our statistical reports are based on statistical surveys which contain both nonsampling errors and sampling errors. While the sampling errors are calculable, the nonsampling errors are unknown, though we know they are substantial in some circumstances, particularly at the state or lower levels of aggregation. Our responsibility is to minimize the introduction of error beyond the survey process, and to the extent possible help survey directors identify sources of error, particularly nonsampling error in NCES surveys.

Despite reasonable efforts on our part, several notable errors have cropped up in program reports. These errors do not appear to be systematic, but instead result from different kinds of human error, so that a simple solution, such as editing code for a particular program, does not seem to be a likely solution. Errors have been a significant problem in our publication process for as long as I have been involved in statistical publications, about 17 years.

My first assignment at NCES was fact checking *Statistics of State School Systems, 1975-76*. I found a significant number of errors. Some of which had been continued for nearly 20 years. At that time publications were largely produced by hand, frequently by senior level statisticians. Trend tables were prepared by cutting out the previous years table, drawing in the columns for the current year, posting the current year by long hand, and having the whole thing set by typesetter, then checked. As a result, the data were very carefully reviewed because of all of the hand work, and big errors were pretty rare. But, small mistakes were very common, much more so than today. Moreover, errors

crept in over time and were compounded over the years A couple of errors in the body of big state trend tables was pretty common. Generally, these were typos or transcription errors compounded by the number of data points on the table. Today, the situation is somewhat different. We are producing vastly more per staff member than we were at that time. Also, the staff are generally more junior level. The extensive use of contractor staff serves to exacerbate this situation. Extensive application of programs like Lotus has made us much more productive and has dramatically reduced to the potential for errors and has greatly eased the correction process. Another important factor is that each trend table is produced from an original file so that errors do not get compounded to the same extent.

The new situation is that there is less individual review of materials since each analyst is producing much more than formerly. The potential now is for fewer errors in terms of a rate, but potentially bigger ones, like entire columns labeled wrong or rows of data loaded into the wrong portion of a table, or incorrect totals entered on critical tables. We have very few transcription errors because so much of the work is done by computer, but computers enable really massive screw ups because of the volume of data being processed.

Optimally, we should have no (non survey) errors in our publications. This does not really seem really feasible because of the staff and resource requirements. The only way to do this would be the traditional "Key and Verify" approach which would require two separate groups to produce the *Digest, Condition, Projections,* etc. and to have a computer cross check. Then the discrepancies would have to be examined, most of which would be various types of format differences. This would require about triple the current resources. Other program groups would face similar problems.

More realistically, we are faced with a situation where no additional resources are available and no additional time is allowed in the production process. The *Condition* ended up nearly 2 months late this year, despite my best efforts to push it forward quickly. Obviously, the absence of Tom Smith was partly a factor, but the main point is that we have to think about speeding up the process rather than slowing it down. Other publications like the *Digest, Projections, and Youth Indicators,* have been delayed, though some of this is a result of the furloughs.

In our examination and resolution of error problems, we need to strike a careful balance between resources and accuracy requirements.

Types of Error

Most errors in charts and tables arise from posting errors, programming errors, translation errors, and labeling errors.

- Posting errors arise when a number is incorrectly entered into a table, either in the table construction phase or the type setting phase. For an example, the recent error in the *Projections* publication on the state total was apparently a posting error in the typesetting process. These errors occur on a random basis, the frequency varying upon the care and skill of the analyst.

- Programming errors are a result of improper manipulations of data. In the dropout error in the *Digest* this was an error on a Lotus spreadsheet that excluded the dropout statistics for Hispanic women. This errors are likely a major problem with Center products. By and large, tabulations are only checked for internal consistency. There are few cases where tabulations are cross-checked by different programmers. One of the few cases where this work is done is for the *Digest*. We routinely compare the output of the IPEDS staff with material for the *Digest*. Also, the Digest uses many of the same SAS programs every year which have been verified many times. This minimizes programming error. However, many manipulations are made through Lotus and involve humans, ergo the likelihood of some error.

- Labeling errors may sound innocuous, but in fact are major sources of error. This type of problem is probably less common in the *Digest* or *Projections* than other NCES publications such as the *Condition*, which have brand new tabulations with each new edition. In the *Digest*, the most common labeling problem is when the analyst updates a trend table and forgets to add the new year label and just copies the old year. The *Digest* also benefits from retaining most of the same tables from year to year. Over time, labeling inconsistencies and misleading labels tended to be weeded out.

- Misleading labels are another related problem. Labels need to be precise, but also concise. Labels need to tell people what the variable is, but they need to be comprehensible. Sometimes we make titles and labels so long and complex that people get confused about what something is and make their own interpretations. We may comfort ourselves that we have "clarified" something with caveats, but if most people are misinterpreting a label, we have something closely akin to an error. My preference is to have a fairly simple label that makes at least the genre of the item clear, and add technical information and clarification in a footnote. The best way to clear up labeling problems is to have a relatively large number of reviewers, which should expressly include people who are not directly familiar with the data under investigation. It is very common for labels to be obvious to the survey directors and obscure to non-initiates. The Condition poses particular problems in this regard, because it presents technical information that is complex, yet (hopefully) designed for more general audiences

Action Plan for Reducing Error

While we have all accepted the need for tighter quality control, we need to bear in mind that additional resources, in time, material, and staff, are very limited. Because of the varying nature of the *Digest, Condition,* and *Projections,* I will suggest somewhat different approaches to focus on the type of errors most likely to occur with each publication.

General System and Organization

The reorganization of the *Condition, Digest,* and Projections into one work unit a few months ago is a improvement that enables closer working relationships. Now, when tabulations are generated for the Condition, we can easily produce the related *Digest* tables at the same time. This will improve operational efficiency and give us more time to work on quality control. Also, *Digest* and *Projections* use the same historical databases. These can be more easily integrated, like they were before the *Projections* team was broken off from the *Digest* team some years ago. Similarly, the Condition team was in the same branch as the *Digest* years ago.

Some of the *Condition* tables may be arranged to draw more directly on *Digest* tables or even computer linked to *Digest* databases/tables to extract data. This will result in more comparable data from publication to publication and also reduce possibilities for error, as well as speed up the process.

Digest Specific Actions

Programming error—This type of error is not so much a problem with annual databases since we use SAS coding that has been tested over the years. However, problems can arise in manipulations after the data have been downloaded. We will add diagnostic check totals to the table program logic that will enable us to check the totals against final tables. This type of check exists for some programs such as the NCES-developed IPEDS programs, but not for CPS dropouts statistics. The already extensive use of databases and automated extraction macros serves to reduce this type of problem. Errors in databases have been easy to locate since they produce rampant errors that are easily detectable. We will increase the use of databases for additional tables to further reduce posting and programming (by table manipulation) errors.

Labeling errors—These tend to be less of a problem with the *Digest* than other publications, but still pose as significant sources of error. The most common ones are

failure to update all of the dates in a table. These are pretty easy to spot. We will institute a cross check of the final year of data in a table with the date printed on the table title. This will occur after the table has been typeset, since these errors are sometimes created during the typesetting process. New tables which are by far most prone to labeling errors will be identified for review by Mary Frase.

In spite of specific actions, some errors will occur because staff continue to be composed of humans. Accepting that some error is likely to be inevitable without unrealistically massive additional resources, we need to reduce the impact of remaining errors. I recommend that we impose an additional review of key tables before the publication goes to print, after ***all*** changes and revisions have been made. This last-ditch review will be designed to avoid embarrassing errors on tables that are used frequently.

The checks are as follows:

Tables 1, 2, 3 cross check enrollment with Tables 38, 39, 167, and 184

Table 4 cross check with tables 63, 64, and 217

Table 8 cross check with table 9

Table 81 cross check with tables 82 and 85

Table 91 check column labels and title

Table 101 and 102 internal check

Table 108 check against NAEP reports

Table 115 check against NAEP reports

Table 122 check against NAEP reports

Table 135 check

Tables 157, 158, 159, 160, and 161 cross check

Table 164 - cross check (using appropriate adjustment) with total on 165; check Wyoming data internally.

Table 172 cross check with tables 181, 182, 183 and 184

Table 210 check table totals and labels with data.

Table 236 cross check with table 241

Table 241 cross check with table 243 and 244

Table 318 cross check with table 321, 322, and 323

Table 328 cross check with table 329, 333, and 334

Table 351 cross check totals with 352 and 353

Table 375 cross check with table 177 and 178

Most of the cross checks can be accomplished very expeditiously. This work should be easily accomplished in a couple of days. The tables are selected based on those that are widely used, those that have been prone to error in the past, and those that are developed in such as way as to check system inputs.

The recommendations above can be accomplished with existing resources and will result in only modest delays for publications. Some of the new procedures will be investments that will save time in the future.

Projections Specific Actions

I am still developing this section. I am planing an automated cross check of all of the enrollment tables in the publication, which will incorporate about 2/3 of the tables. This will not help guard against typesetting errors. I will work with *Projections* staff to develop additional ideas.

Condition Specific Actions

These are most problematic of all, since I think most of the errors are of the programming variety which are most difficult to find. Also, the time lines for the publication is so tight that little time for review is available. I have a couple of thoughts that should be discussed more broadly because they affect the nature of the publication.

We should attempt to use more off-the-shelf materials that have data or totals that correspond with other publications, such as the *Digest* for cross checking purposes. We need to retain the specialized nature of some of the tables, but we should use standard

NCES tabulations or definitions when we can. Consideration of this policy has been urged by the Chief Statistician.

We should also try to avoid pushing the envelop on release of data. Many data items come in so late that there is little time for review or they get revised during the group review. In some cases, we have had to revise indictors at the last minute because data were not released in time.

Consider standardizing a few more of the tables so that so many new tabulations are not required every year.

Consider not using color for charts. The use of color results in considerable delays and problems in printing when the resources might be more effectively used for review tables and ensuring up-to-date figures. In any case, we need to develop a much better procedure for printing the publication in the future. It took almost 2 ½ months to get this year's edition printed after it was "completely" type set.

BEST COPY AVAILABLE

Appendix B. Application of Recommendations to Previously Identified Problems

The following is a summary of the four problem areas which originally led to the creation of this Task Force. The list includes a description of the source of error and cites specific Task Force recommendations for avoiding similar situations in the future.

Report	Type of Error	Source of Error	Recommendation*
Projections in Education Statistics	Incorrect enrollment totals on tables	Typesetting	• Perform checks on camera-ready copy similar to those identified in *Digest* memo (see Appendix A). • Use benchmarks to verify totals of selected tables (#10). • Generate supplementary cross-tabulations (#11).
Projections of Education Statistics	Incorrect projections for two states	Incorrect keying of data	• Incorporate check totals for data to be used in projections model (#16).
Digest of Education Statistics	Incorrect figure in dropout table	Incorrect manipulation of formula calculations in spreadsheets	• Double check selected tables after the typesetting process for final verification of critical items. • Include input from survey staff when doing secondary analysis (#3). • Improve documentation of calculations performed to produce figures (#9). • Check for consistency across publications (#7).
Allocations	Incorrect figure for Maine	Failure to integrate all phases of legislation into formula	• As per determination by Commissioner, an additional staff member is to be added to the project for more review steps (#15).

* The numbers in parentheses refer to Task Force recommendations in Ch.II, Section 2.

Appendix C. References

Booth, W.C., Colomb, G.G., and Williams, J.M. 1995. *The craft of research*. Chicago: University of Chicago Press.

Federal Committee on Statistical Methodology. 1991. *Seminar on quality of Federal data: Statistical Policy Working Paper 20, (Chapter II of 3)*. Washington, DC: Office of Management and Budget. (NTIS No. PB91-142414).

Joint Economic Committee. 1991. *Improving Federal Statistics*. Hearing before the Joint Economic Committee, Congress of the United States: One hundred second Congress, first session. (S. Hrg. 102–184). Washington, DC: U.S. Government Printing Office.

Miller, D.C. 1991. *Handbook of research design and social measurement* (5th ed.). Newbury Park, CA: Sage Publications.

Nagel, S. S., and Neef, M. 1979. *Policy analysis in social research*. Beverly Hills: Sage Publications.

Sundgren, B. and Dean, P. 1996. *Metadata: A quality element in official statistics — the Swedish approach*. Statistics Sweden.

Wolins, L. 1982. *Research mistakes in the social and behavioral sciences*. Ames, IA: Iowa State University Press.

Listing of NCES Working Papers to Date

Please contact Ruth R. Harris at (202) 219-1831
if you are interested in any of the following papers

Number	Title	Contact
94-01 (July)	Schools and Staffing Survey (SASS) Papers Presented at Meetings of the American Statistical Association	Dan Kasprzyk
94-02 (July)	Generalized Variance Estimate for Schools and Staffing Survey (SASS)	Dan Kasprzyk
94-03 (July)	1991 Schools and Staffing Survey (SASS) Reinterview Response Variance Report	Dan Kasprzyk
94-04 (July)	The Accuracy of Teachers' Self-reports on their Postsecondary Education: Teacher Transcript Study, Schools and Staffing Survey	Dan Kasprzyk
94-05 (July)	Cost-of-Education Differentials Across the States	William Fowler
94-06 (July)	Six Papers on Teachers from the 1990-91 Schools and Staffing Survey and Other Related Surveys	Dan Kasprzyk
94-07 (Nov.)	Data Comparability and Public Policy: New Interest in Public Library Data Papers Presented at Meetings of the American Statistical Association	Carrol Kindel
95-01 (Jan.)	Schools and Staffing Survey: 1994 Papers Presented at the 1994 Meeting of the American Statistical Association	Dan Kasprzyk
95-02 (Jan.)	QED Estimates of the 1990-91 Schools and Staffing Survey: Deriving and Comparing QED School Estimates with CCD Estimates	Dan Kasprzyk
95-03 (Jan.)	Schools and Staffing Survey: 1990-91 SASS Cross-Questionnaire Analysis	Dan Kasprzyk
95-04 (Jan.)	National Education Longitudinal Study of 1988: Second Follow-up Questionnaire Content Areas and Research Issues	Jeffrey Owings
95-05 (Jan.)	National Education Longitudinal Study of 1988: Conducting Trend Analyses of NLS-72, HS&B, and NELS:88 Seniors	Jeffrey Owings

Listing of NCES Working Papers to Date--Continued

Number	Title	Contact
95-06 (Jan.)	National Education Longitudinal Study of 1988: Conducting Cross-Cohort Comparisons Using HS&B, NAEP, and NELS:88 Academic Transcript Data	Jeffrey Owings
95-07 (Jan.)	National Education Longitudinal Study of 1988: Conducting Trend Analyses HS&B and NELS:88 Sophomore Cohort Dropouts	Jeffrey Owings
95-08 (Feb.)	CCD Adjustment to the 1990-91 SASS: A Comparison of Estimates	Dan Kasprzyk
95-09 (Feb.)	The Results of the 1993 Teacher List Validation Study (TLVS)	Dan Kasprzyk
95-10 (Feb.)	The Results of the 1991-92 Teacher Follow-up Survey (TFS) Reinterview and Extensive Reconciliation	Dan Kasprzyk
95-11 (Mar.)	Measuring Instruction, Curriculum Content, and Instructional Resources: The Status of Recent Work	Sharon Bobbitt & John Ralph
95-12 (Mar.)	Rural Education Data User's Guide	Samuel Peng
95-13 (Mar.)	Assessing Students with Disabilities and Limited English Proficiency	James Houser
95-14 (Mar.)	Empirical Evaluation of Social, Psychological, & Educational Construct Variables Used in NCES Surveys	Samuel Peng
95-15 (Apr.)	Classroom Instructional Processes: A Review of Existing Measurement Approaches and Their Applicability for the Teacher Follow-up Survey	Sharon Bobbitt
95-16 (Apr.)	Intersurvey Consistency in NCES Private School Surveys	Steven Kaufman
95-17 (May)	Estimates of Expenditures for Private K-12 Schools	Stephen Broughman
95-18 (Nov.)	An Agenda for Research on Teachers and Schools: Revisiting NCES' Schools and Staffing Survey	Dan Kasprzyk
96-01 (Jan.)	Methodological Issues in the Study of Teachers' Careers: Critical Features of a Truly Longitudinal Study	Dan Kasprzyk

Listing of NCES Working Papers to Date--Continued

Number	Title	Contact
96-02 (Feb.)	Schools and Staffing Survey (SASS): 1995 Selected papers presented at the 1995 Meeting of the American Statistical Association	Dan Kasprzyk
96-03 (Feb.)	National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues	Jeffrey Owings
96-04 (Feb.)	Census Mapping Project/School District Data Book	Tai Phan
96-05 (Feb.)	Cognitive Research on the Teacher Listing Form for the Schools and Staffing Survey	Dan Kasprzyk
96-06 (Mar.)	The Schools and Staffing Survey (SASS) for 1998-99: Design Recommendations to Inform Broad Education Policy	Dan Kasprzyk
96-07 (Mar.)	Should SASS Measure Instructional Processes and Teacher Effectiveness?	Dan Kasprzyk
96-08 (Apr.)	How Accurate are Teacher Judgments of Students' Academic Performance?	Jerry West
96-09 (Apr.)	Making Data Relevant for Policy Discussions: Redesigning the School Administrator Questionnaire for the 1998-99 SASS	Dan Kasprzyk
96-10 (Apr.)	1998-99 Schools and Staffing Survey: Issues Related to Survey Depth	Dan Kasprzyk
96-11 (June)	Towards an Organizational Database on America's Schools: A Proposal for the Future of SASS, with comments on School Reform, Governance, and Finance	Dan Kasprzyk
96-12 (June)	Predictors of Retention, Transfer, and Attrition of Special and General Education Teachers: Data from the 1989 Teacher Followup Survey	Dan Kasprzyk
96-13 (June)	Estimation of Response Bias in the NHES:95 Adult Education Survey	Steven Kaufman
96-14 (June)	The 1995 National Household Education Survey: Reinterview Results for the Adult Education Component	Steven Kaufman

Listing of NCES Working Papers to Date--Continued

Number	Title	Contact
96-15 (June)	Nested Structures: District-Level Data in the Schools and Staffing Survey	Dan Kasprzyk
96-16 (June)	Strategies for Collecting Finance Data from Private Schools	Stephen Broughman
96-17 (July)	National Postsecondary Student Aid Study: 1996 Field Test Methodology Report	Andrew G. Malizio
96-18 (Aug.)	Assessment of Social Competence, Adaptive Behaviors, and Approaches to Learning with Young Children	Jerry West
96-19 (Oct.)	Assessment and Analysis of School-Level Expenditures	William Fowler
96-20 (Oct.)	1991 National Household Education Survey (NHES:91) Questionnaires: Screener, Early Childhood Education, and Adult Education	Kathryn Chandler
96-21 (Oct.)	1993 National Household Education Survey (NHES:93) Questionnaires: Screener, School Readiness, and School Safety and Discipline	Kathryn Chandler
96-22 (Oct.)	1995 National Household Education Survey (NHES:95) Questionnaires: Screener, Early Childhood Program Participation, and Adult Education	Kathryn Chandler
96-23 (Oct.)	Linking Student Data to SASS: Why, When, How	Dan Kasprzyk
96-24 (Oct.)	National Assessments of Teacher Quality	Dan Kasprzyk
96-25 (Oct.)	Measures of Inservice Professional Development: Suggested Items for the 1998-1999 Schools and Staffing Survey	Dan Kasprzyk
96-26 (Nov.)	Improving the Coverage of Private Elementary-Secondary Schools	Steven Kaufman
96-27 (Nov.)	Intersurvey Consistency in NCES Private School Surveys for 1993-94	Steven Kaufman

Listing of NCES Working Papers to Date--Continued

Number	Title	Contact
96-28 (Nov.)	Student Learning, Teaching Quality, and Professional Development: Theoretical Linkages, Current Measurement, and Recommendations for Future Data Collection	Mary Rollefson
96-29 (Nov.)	Undercoverage Bias in Estimates of Characteristics of Adults and 0- to 2-Year-Olds in the 1995 National Household Education Survey (NHES:95)	Kathryn Chandler
96-30 (Dec.)	Comparison of Estimates from the 1995 National Household Education Survey (NHES:95)	Kathryn Chandler
97-01 (Feb.)	Selected Papers on Education Surveys: Papers Presented at the 1996 Meeting of the American Statistical Association	Dan Kasprzyk
97-02 (Feb.)	Telephone Coverage Bias and Recorded Interviews in the 1993 National Household Education Survey (NHES:93)	Kathryn Chandler
97-03 (Feb.)	1991 and 1995 National Household Education Survey Questionnaires: NHES:91 Screener, NHES:91 Adult Education, NHES:95 Basic Screener, and NHES:95 Adult Education	Kathryn Chandler
97-04 (Feb.)	Design, Data Collection, Monitoring, Interview Administration Time, and Data Editing in the 1993 National Household Education Survey (NHES:93)	Kathryn Chandler
97-05 (Feb.)	Unit and Item Response, Weighting, and Imputation Procedures in the 1993 National Household Education Survey (NHES:93)	Kathryn Chandler
97-06 (Feb.)	Unit and Item Response, Weighting, and Imputation Procedures in the 1995 National Household Education Survey (NHES:95)	Kathryn Chandler
97-07 (Mar.)	The Determinants of Per-Pupil Expenditures in Private Elementary and Secondary Schools: An Exploratory Analysis	Stephen Broughman
97-08 (Mar.)	Design, Data Collection, Interview Timing, and Data Editing in the 1995 National Household Education Survey	Kathryn Chandler

BEST COPY AVAILABLE

Listing of NCES Working Papers to Date--Continued

Number	Title	Contact
97-09 (Apr.)	Status of Data on Crime and Violence in Schools: Final Report	Lee Hoffman
97-10 (Apr.)	Report of Cognitive Research on the Public and Private School Teacher Questionnaires for the Schools and Staffing Survey 1993-94 School Year	Dan Kasprzyk
97-11 (Apr.)	International Comparisons of Inservice Professional Development	Dan Kasprzyk
97-12 (Apr.)	Measuring School Reform: Recommendations for Future SASS Data Collection	Mary Rollefson
97-13 (Apr.)	Improving Data Quality in NCES: Database-to-Report Process	Susan Ahmed

BEST COPY AVAILABLE