Proceedings of Statistics Canada Symposium 2001
Achieving Data Quality in a Statistical Agency: A Methodological Perspective

QUALITY ASSURANCE CHALLENGES IN THE UNITED STATES' CENSUS 2000

David C. Whitford and Jennifer W. Reichert, U.S. Census Bureau[1]

ABSTRACT

The United States' Census 2000 operations were more innovative and complicated than ever before. We are required to produce state population totals within nine months and, using our coverage measurement survey, to produce adjusted counts within one year. Therefore, all operations had to be implemented and completed quickly with quality assurance (QA) that was effective and had prompt turnaround. Our QA challenges were to get timely information to supervisors, do prompt checks of "suspect" work, and provide reports to headquarters quickly. This paper will, by presenting these challenges and our solutions, provide an overview of the Census 2000 QA program.

For interviewing operations, we used a random reinterview to check each enumerator's work at the outset and an administrative reinterview program to continuously target enumerators needing attention.

Our reinterview programs used automated selection to ensure accuracy and efficiency. For the random reinterview, the system selected random cases from every enumerator's work for reinterview. For the administrative reinterview, the system compiled informative statistics for each enumerator and used statistical process control methodology to flag enumerators who were "out of control" for those statistics.

In the Census 2000 coverage measurement survey, the reinterview programs were further automated by making use of a Computer Assisted Personal Interviewing (CAPI) environment. With one-day turnaround, supervisors received QA summary data on their laptop computers.

After interviewing, we matched the coverage measurement interviews with the census interviews. In Census 2000, this operation was quick and paperless. We specified a completely automated QA system to derive the operation's outgoing quality.

Finally, all census data were captured using imaging, Optical Mark Recognition (OMR), Optical Character Recognition (OCR), and keying. We contracted out these tasks and developed a QA program that would ensure an acceptable end product of the automated data capture system. Our system monitored a sample of the images and data from the contractor to determine if the captured data were accurate.

KEY WORDS: Quality Assurance; Reinterview; Coverage Measurement; Data Capture

1. INTRODUCTION

The United States= Census 2000 operations were more innovative and complicated than those of previous censuses. We made use of many advances in automation, and technical progress introduced the need to hire contractors to undertake specialized duties. The increased automation and contracting presented many new challenges to ensure quality for Census 2000.

The QA philosophy for Census 2000 followed W. Edwards Deming=s approach of preventing defects through process improvement rather than inspection. The objective of process improvement is to build products correctly the first time. Statistical process control plays a major role in achieving this objective as does management involvement and commitment to the quality improvement process. Our QA mission was threefold:

---

[1]This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

- to minimize significant performance errors
- to prevent clustering of significant performance errors, and
- to promote continuous improvement.

With increased utilization of contractors and technology, we faced the challenge of integrating the Deming philosophy and our high QA standards into systems that were either managed by contractor staff or automated so as to limit the amount of human intervention.

One of the reasons for our intensive use of automation and contractors was that the U.S. census must, by law, be produced under a very tight schedule:

- The initial census data had to be captured, processed and compiled by December 31, 2000Bthe deadline for producing the population counts for each of our fifty states.
- By April 1, 2001 we had to complete the detailed census countsBthat is, to allow for the adjustment we needed to apply the results of the coverage measurement survey to the initial census counts at the block level to account for the degree to which people in each block were under- or over-counted. (Ultimately, the Secretary of Commerce decided not to use these adjusted counts for drawing political boundaries; they still may be used as a base for future surveys or for funds distribution.)

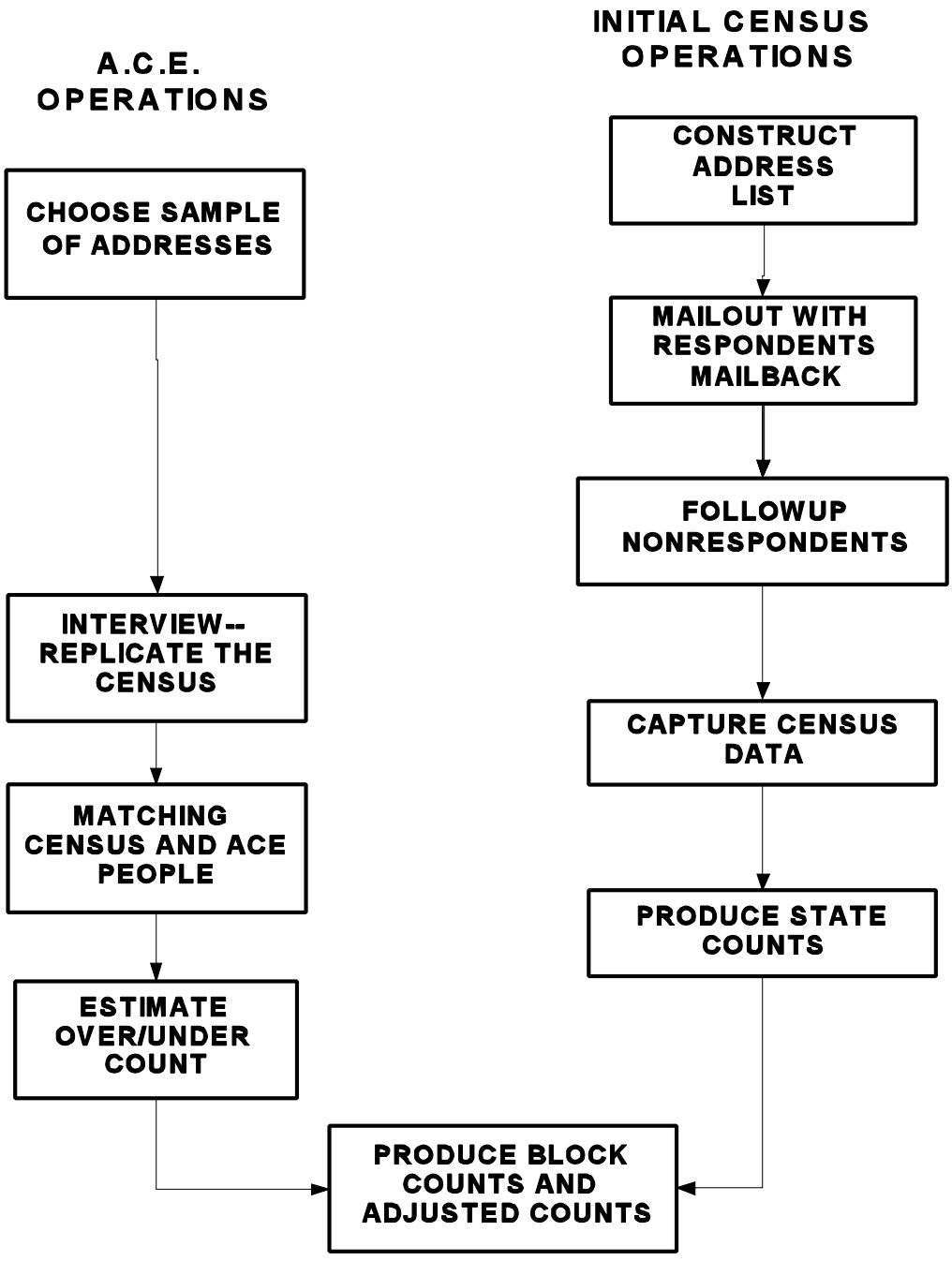
To produce the data for these two deadlines, there were key steps in Census 2000 that we had to complete successfully:

- Build a list of addresses of dwelling units throughout the U.S.,
- Send, drop off, or complete census questionnaires at those addresses,
- Followup with nonrespondents by visiting their addresses, and
- Capture the census data.

Census 2000 also included a coverage measurement survey, known as the Accuracy and Coverage Evaluation (A.C.E.). The A.C.E. survey, too, had to be implemented successfully before the census data could be produced to meet the second deadline above. The basic steps for the A.C.E. survey were to:

- Construct a sampling frameBdividing the country into areas, each of which could be interviewed by a single enumerator.
- Choose an area sample comprising 314,000 dwelling units.
- After the initial steps of the census enumeration were complete, conduct an interview which essentially replicated the census for people living at these addresses.
- Match these results to the initial census: that is, match names and demographic characteristics from the survey and the census to determine census omissions (under-count) and erroneous inclusions (over-count).
- Use these results to estimate how many people and housing units were missed or counted more than once during the initial census steps.

For reference, here is a flowchart of the census operations mentioned thus far:


A.C.E.
OPERATIONS
CHOOSE SAMPLE OF ADDRESSES
INTERVIEW-- REPLICATE THE CENSUS
MATCHING CENSUS AND ACE PEOPLE
ESTIMATE OVER/UNDER COUNT
INITIAL CENSUS
OPERATIONS
CONSTRUCT ADDRESS LIST
MAILOUT WITH RESPONDENTS MAILBACK
FOLLOWUP NONRESPONDENTS
CAPTURE CENSUS DATA
PRODUCE STATE COUNTS
PRODUCE BLOCK COUNTS AND ADJUSTED COUNTS

2. QA CHALLENGES

Census 2000 operations were huge: approximately 400,000 people were employed to followup nonrespondents quickly and approximately another 15,000 to use the highly technical optical scanning devices that quickly captured the data from the census forms. Given this large temporary workforce, the need for a carefully designed QA program became apparent. Tight time constraints and the sheer magnitude of the staff needed to perform census activities combined to hinder our ability to be as selective as we would have liked in recruiting and hiring people to staff the 520 local census offices and three data capture centers. We had to develop adequate and realistic quality and performance expectations prior to designing the QA programs so that the quality checks we implemented would ensure the quality of the census data.

In addition, the Census Bureau applied the latest advances in automation:

- Laptop computers were used for some interviewing operations and supervisory control of other field operations.
- Data were captured from the completed forms using imaging technology and interpreted with optical mark recognition, optical character recognition, and keying.

Some of these technologies were essentially unproven on the huge scale of the census. We again see the need for an extensive QA program. The primary challenges in developing the QA program for the automated processes were to define errors and to develop methods for measuring those errors. Some of the technology was so new, we didn=t have the experience necessary to be able to anticipate all possible errors. This lack of experience significantly increased the difficulty of the error-definition process. Once the errors were defined and measured, we had to design a QA system to detect processes with unacceptable error rates and to correct the errors. Because the purpose of incorporating automation was to limit the amount of human intervention and, hence, human error, the QA systems we designed and implemented had to be as automated as possible too.

To incorporate the latest technology into our operations, the Census Bureau hired contractors whenever their expertise was necessary to facilitate the technology=s implementation. For example,

- Lockheed-Martin was engaged to design and develop the data capture system.
- TRW was responsible for facility acquisition and operation of the data capture centers.
- Gunnison Consulting Group was contracted to further process the results of laptop interviewing in the coverage measurement survey

The Census Bureau was not accustomed to dealing with contractors on such a large scale. We had to adjust our QA approach to deal with these new production entities. The contractors generally implemented their own internal QA programs, but we had to review and evaluate those programs to ensure they met our QA requirements. The true QA challenges associated with the increased contracting were driven by the number of different contractors working on the same operations and the multitude of concurrent processes that each required QA checkpoints.

We will discuss a representative sample of these QA challenges in terms of the census and A.C.E. operations, as follows:

- InterviewingBWe developed innovative QA programs for interviewing both in nonresponse followup and A.C.E. interviewing. These made the best use of our automated resources such that the QA results were distributed quickly and effectively to supervisors.
- MatchingBWe formed a close partnership with the contractor that developed the automated software used to facilitate clerical matching of the census and A.C.E. results. Clerical errors in matching were

difficult to define and discern. A very important topic of interest with respect to deciding on the eventual usability of the survey was the computation of outgoing error.

$ Data CaptureBLockheed-Martin created a system to capture data from the approximately 148 million forms received from respondents and enumerators. The data capture system (DCS2000) was the most technologically advanced system ever used in a decennial census. The QA challenge of monitoring the quality of the data coming out of the DCS2000 involved developing independent monitoring procedures to match the technological sophistication of the contractor=s system.

3. INTERVIEWING

Enumerators contacted households to conduct interviews in both the enumeration phase of Census 2000 and during the coverage measurement phase (A.C.E.). They contacted over 42 million households during enumeration and over 300,000 during the A.C.E. Census 2000 was the largest interviewing operation the Census Bureau has ever undertaken.

The quality assurance systems employed in the 1990 census relied mostly on shuffling paper. That is, completed census (or coverage measurement survey) questionnaires would be collected at a central office, a QA sample of them would be selected for reinterviewing, addresses would be transcribed from the questionnaire to the QA form, the QA interview would be conducted, its pass or fail status determined, and word of failures passed on to the original interviewer=s supervisor. Obviously, it could take this system a couple of weeks to detect problem interviewers. And most of the interviewing operations outlined above took six weeks or two months in their entirety.

Our challenge for Census 2000 was to automate these activities and complete them within a more effective time frame.

For all of the interviewing conducted by census enumerators, the Census Bureau implemented a reinterview program to ensure the quality of the data collected during the interviews. The goal of the reinterview program was to detect and deter discrepant results. The reinterview program acted as a deterrent to interviewer falsification by providing early, up-front prevention of errors, both intentional and unintentional. To provide long-term protection, the reinterview program was a successful device for detecting when enumerators were submitting faulty data. The reinterview programs consisted of administrative and random components:

- Administrative – The administrative reinterview program identified enumerators for reinterview by collecting data from enumerators' completed work, producing statistics from those data, and producing reports for supervisors that compared characteristics of their work to that of the other enumerators in their enumeration area. Any enumerators whose work was significantly different were flagged on the reports and more of their cases possibly placed into reinterview by their supervisors. The administrative reinterview was a targeted approach that identified suspect work quickly and provided a higher probability of detection.

- Random – The random reinterview program was implemented by a computer system that selected a random sample of cases from the initial cases that every enumerator completed to be placed into reinterview. This type of reinterview provided for broad protection, and thus an effective deterrent, across all enumerators.

These two reinterview components complimented each other to provide comprehensive QA coverage of the interviewing activities.

All field operations for Census 2000 had to be completed within three months. For that reason, it was imperative that the supervisors received the results of the QA operations promptly so they had time to review

the results and take appropriate actions. In order for this quick turnaround of QA results to happen, the process for selecting, distributing, and checking reinterview cases had to be extremely efficient.

To ensure efficiency in the reinterview operations, we implemented automated procedures wherever possible. The selection of cases for reinterview was automated for both the administrative and random reinterview programs. The ACE reinterviewers used computer assisted personal interviewing (CAPI) for their reinterviews.

A specially-designed computer system automatically selected cases for reinterview for all enumerators placed into the reinterview system. In the case of the random reinterview, the system selected a random sample of cases from every enumerator=s workload. For the administrative reinterview system, reinterview supervisors would review the results of the administrative test and identify enumerators to be reinterviewed. The system would then select cases from those enumerators= workloads to be reinterviewed.

For the A.C.E., the reinterview operation was completely automated using the selection procedures described above as well as CAPI. The reinterview enumerators conducted their reinterviews by entering responses directly into an electronic questionnaire on laptop computers. For the other reinterview operations, the enumerators recorded respondents= responses on paper forms. The CAPI system provided for nearly immediate transfer of information from completed interviews to the reinterview supervisors and, in turn, back to the enumerators. Reinterview cases were transmitted to the reinterviewers= laptops and the results of the reinterviews were transmitted back to the supervisors on a daily basis. This level of automation allowed the supervisors to perform timely reviews so they could take appropriate action with little or no delay.

For the other reinterview operations that used paper questionnaires, the Census Bureau had to ensure adequate staffing to get the questionnaires through the reinterview process as quickly as possible. The reinterview staffing consisted primarily of an Office Operations Supervisor (OOS), Reinterview Crew Leaders (RCL), and reinterviewers. The RCLs were responsible for distributing the reinterview cases to the reinterviewers and also for collecting the completed reinterview cases and transporting them to the Local Census Office (LCO). In the office, the OOS would review the results of the reinterview and recommend appropriate actions to the Assistant Manager for Field Operations and the LCO Manager. Because some of the enumeration areas could be a significant distance from the local offices, the shuttling of forms to and from the field presented a significant challenge for the reinterview staff.

4. ACCURACY AND COVERAGE EVALUATION MATCHING

Subsequent to the census interviewing activities, the next step in the coverage measurement survey processing was to match the census and the A.C.E. interview results. Matching gave us an estimate of the number of people missed in the census and the number of people erroneously enumerated.

The 1990 version of clerical matching was spread out over seven processing offices. Although a great amount of effort was put into maintaining between-site consistency, our evaluation of the results showed that some processing office effects had crept into the system. We had to avoid this in 2000. In Census 2000 we first matched census and A.C.E. by computer and then had clerks solve the cases which were unresolved by the computer match. For the clerical match we employed a contractor to develop computerized clerical matching software that allowed the clerk to examine, and hopefully match, the unresolved household records from the census and the A.C.E.

The system allowed clerical matching to be an automated, Apaperless@ activityBquite an advance from previous censuses where clerks were required to retrieve match forms, census and A.C.E. records, address lists, and maps before they could begin their matching work. The automated system cut the production staff we needed by close to a factor of ten from what it was in the 1990 census coverage measurement survey. This, of course, allowed for much tighter control of the quality of matching.

There were 250 clerks employed to match the unresolved records. Since their work was so important, 50 matching technicians did the QA of the clerks and handled the tough cases the clerks referred to them. Finally, 12 analysts were employed to handle the very hardest cases and perform the QA of the technicians. In this way, the most difficult cases were effectively quality controlled. Analysts, in this case full time Census Bureau employees who have done this work for years, were assumed to be error-free.

Additionally, the automated clerical matching system could detect whole work units of cases that needed to be completely reviewedBfor instance, an analyst reviewed all work units where a technician, during QA, changed over 50 percent of a clerk's work.

In the remaining cases, we imposed a stringent, automated QA plan.

When the A.C.E. matching task began, all clerks and technicians had 100 percent of their work checked. The results of the check were automatically recorded. Once 200 of their codes had been checked and their error rate was less than four percent, only a sample of their subsequent work was checked. If a clerk's or technician's error rate rose above four percent, their work was once again subject to a 100 percent check.

The computerized matching system allowed headquarters employees to monitor the entire system. In addition, the system provided headquarter=s management with the ability to change the QA parametersBfor instance, changing the four percent cutoff mentioned above to a higher/lower number, if necessary. Since, as mentioned above, the A.C.E. had to be finished under a tight schedule, managers needed this option if the work was moving too slowly through any given matching stage.

One challenge of this entirely automated system was computing the outgoing qualityBa statistic of interest to Census 2000 oversight and advisory entities (e.g., the U.S. Congress, and the U.S. National Academy of Sciences). The calculation had to allow for the extraction of Areview@ records from work units and for the sampling rate of clerks being switched from sample to 100 percent and back again. However, the matching system saved the history of every code and, although cumbersome, we derived adequate outgoing quality measures.

The paperless, computerized clerical matching system and the QA system allowed for the matching activity to be very tightly controlledBone instance in which contracting and automation were very helpful in allowing us achieve a quality census.

5. DATA CAPTURE

Because nearly all of the data collected during Census 2000 was recorded on paper questionnaires, the Census Bureau had to develop a system to accurately and efficiently extract the data from the questionnaires into an electronic format to allow for data processing and analysis. The traditional method for this extraction is to have keyers key the data directly from the questionnaire into a data file. For Census 2000, the questionnaires were electronically scanned to create image files. Optical Character Recognition (OCR) software interpreted the handwritten responses and Optical Mark Recognition (OMR) software interpreted the check-box responses.

The OCR and OMR software captured all data from the questionnaires, but also assigned a score to indicate the confidence with which the data were captured accurately. For those data items that were below the acceptable confidence score, keyers keyed the data using a Key from Image system. This data capture technology allowed real-time capture of respondent data and prompt availability of the census data.

The Census Bureau contracted out the immense and highly technical data capture activities. All questionnaires were either mailed directly from respondents or shipped by Federal Express from the LCOs to one of four Data Capture Centers. Approximately 148 million forms were scanned into images. Due to the importance of the

data from the census questionnaires, it was imperative that the data capture be an accurate process. Our challenge was to ensure that the contractors achieved that accuracy.

The contractors implemented QA on both the OMR/OCR capture and the keying results. However, the contractors' QA plans focused on the level of accuracy of the census data overall. The Census Bureau was also interested in monitoring the accuracy of the census data at the field level. That is, we wanted to monitor the captured data to identify any large-scale or systematic errors for particular questionnaire items. We established a separate contract to develop an independent QA system to monitor the accuracy of the field-level data.

We developed a system that enabled analysts to view the questionnaire images and the captured data values. The analysts would compare the captured data and the images to identify and categorize any discrepancies. We pre-selected a sample of 768,000 forms -- approximately 0.5 percent of the total forms scanned. The analysts performed the independent QA monitoring on this sample. Due to the speed of the data-capture process, we faced significant challenges in getting the images and data reviewed promptly enough to be able to address and rectify any errors.

6. CONCLUSIONS

Overall, our approach to the quality assurance program employed in Census 2000 made monumental strides over our approach in 1990. However, it is apparent that the census itself also made huge advances in automation, the use of the latest technology, and the employment of specialized contractors who could implement these technologies. The question then is whether the QA advances kept up with the technical advances.

We are attempting to answer this question with an evaluation of our QA approach. We have contracted an outside quality assurance specialist organization that will review the census operations, our QA approach to them, and the debriefings of the staff that implemented the QA. Additionally, the experts will interview Census Bureau staff involved with the implementation of these programs. We hope that this study will answer our questions and provide us with a basis upon which to build for the census of 2010.

On the surface, though, the QA program in Census 2000 seems to have done its job well. We have good solid evidence already that the results from representative examples presented here were acceptable. Interviewing was completed on time with consistent QA results evidencing themselves from all across the country. Our coverage measurement matching system had an exceptional outgoing quality that was confirmed by our oversight entities. And our data capture system has loudly been proclaimed as being successful.

We experienced many challenges in developing a quality assurance program for Census 2000. The challenges were met with innovation and tenacity—two elements that quality products depend upon.