UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001

May 22, 2012

2010 CENSUS PLANNING MEMORANDA SERIES

No. 195

MEMORANDUM FOR The Distribution List

From: Arnold Jackson *[signed]*
Acting Chief, Decennial Management Division

Subject: 2010 Census Decennial Response Integration System Paper Questionnaire Data Capture Assessment Report

Attached is the 2010 Census Decennial Response Integration System (DRIS) Paper Questionnaire Data Capture Assessment Report. The Quality Process for the 2010 Census Test Evaluations, Experiments, and Assessments was applied to the methodology development and review process. The report is sound and appropriate for completeness and accuracy.

If you have questions about this report, please contact Dave Coon at (301) 763-9377.

Attachment

May 21, 2012

2010 Census
Decennial Response Integration System
Paper Questionnaire Data Capture
Assessment Report

U.S. Census Bureau standards and quality process procedures were applied throughout the creation of this report. Census Bureau Management has authorized release of all data contained in this report.

Final

David Coon
Neil Osborne

Decennial Management Division

Raymond Muenzer

Decennial Automated Contract Management Office

Table of Contents

LIST OF FIGURES AND TABLES

Executive Summary

The 2010 Census Decennial Response Integration System (DRIS) Paper Questionnaire Data Capture operation were responsible for the capture and conversion of data from paper questionnaires that were mailed back by respondents, Be Counted and paper-based field operations, as well as paper questionnaire inputs from telephone-based questionnaire operations. The scope of this assessment report is primarily limited to paper questionnaires data captured by DRIS, but the report also includes some high level discussion of the data integration components of DRIS.

For the 2010 Census, there were two contracted data capture centers. The contract centers were located in Phoenix, Arizona, and Baltimore, Maryland. In addition, the National Processing Center in Jeffersonville, Indiana served as a third data capture center, staffed by Federal government personnel.

Together, the three data capture centers processed and captured data from over 164 million paper questionnaires during the 2010 Census. The 164 million questionnaires contained over 3 billion individual checkbox fields. In addition to the 164 million questionnaires processed in calendar year 2010, for Group Quarters Validation operations in 2009, DRIS processed just over two million forms at the National Processing Center.

The decision to have a paper-based Nonresponse Followup census methodology added over 63.5 million enumerator paper questionnaires to the workload for data capture. These forms are included in the 164 million questionnaires figure above. The data capture centers handled the additional processing requirements successfully. Through change requests, this change to a paper-based Nonresponse Followup added over $161 million in contract value to the DRIS program. DRIS was designed with scalable cluster-based architecture. Had DRIS not used scalable architecture, this change could have cost much more to implement.

Because of capacity problems with the Paper Based Operations Control System, Census Bureau management made a decision to remove shipping functionality from The Paper Based Operations Control System. Local Census Offices were unable to ship materials for the first several days of Nonresponse Followup, and materials were accumulating in the Local Census Offices. The removal of shipping functionality from the Paper Based Operations Control System broke significant reconciliation features of DRIS. This resulted in additional staffing expense and end of program reconciliation costs, but DRIS was able to handle the system interface problems during operations. Overall, DRIS had more than fifty different interfaces (system data exchanges) with various Headquarters systems during peak production.

Overall, for mail return forms, the mail return workloads processed by DRIS were below initial projections due to contingency and changes to the replacement mailing strategy that were not in the initial workload estimates.

- The Paper data capture portion of the contract has estimated costs totaling $473,642,000, including paper operations costs of $248,714,000.

- The Workflow Control and Management (data integration) estimated costs totaled $23,749,000.

These numbers include the cost of paper data capture and the proportional cost of Project Management, Engineering, Architecture, Test, Data Quality, Security, Operation Management, contractor procured test materials, Telecommunications, Workflow Control and Management and Electronic Suitability Assessment (fingerprinting) costs. The cost of an individual segment cannot be used to represent the cost of a stand-alone paper questionnaire data capture contract. Cost savings were realized by integrating these segments into one contract. Taking the total paper form cost of $473,642,000 (this includes labor, systems, development), and dividing by the total number of paper forms processed by DRIS, yields an average data capture cost of approximately $2.94 per paper form processed through data capture in DRIS.

DRIS provided all contract deliverables in a timely manner. The DRIS program ended successfully, under budget, and on schedule. DRIS met all contractual requirements.

1. Introduction

Scope

The 2010 Census Decennial Response Integration System (DRIS) Paper Questionnaire Data Capture Assessment Report documents what happened during the 2010 Census DRIS Paper Questionnaire Data Capture operation. This assessment will provide data for the next planning cycle for the 2020 Census and for other Census Bureau operations.

The scope of this assessment covers the paper data capture performed by DRIS and will briefly touch on data integration at a high level as it relates to paper data capture performed by DRIS. DRIS Telephone Questionnaire Assistance, DRIS Coverage Followup, and DRIS Paper Production Data Quality are covered separately by other assessments.

This assessment includes key lessons learned and recommendations and it will document final workloads, costs and lessons learned for all aspects of the DRIS Paper Questionnaire Data Capture operation, including Cost and Progress reports.

Data Quality metrics and analysis related to DRIS are only discussed at a high level within this assessment specifically as they relate to the research questions concerning Reprocessing and Field Operation's impact on data capture.

Intended Audience

This document assumes that the reader has at least a basic understanding of DRIS paper data capture and data integration activities. The goal is to use this document to help research, planning, and development teams planning the 2020 Census. If you do not have a basic understanding of DRIS, please refer to the Census 2010 Informational Memorandum No. 60, the 2010 Census Detailed Operations System Plan (DOSP) for the Data Capture and Integration Operation. The DOSP is a document that describes the overall Data Capture and Integration process (of which DRIS is only one component) in more detail.

The purpose of this assessment is to document the results and major findings from the 2010 Census DRIS paper questionnaire data capture and integration, including topics such as workload, staffing, schedule, and cost. This assessment will inform stakeholders, and decision makers of recommended changes or improvements for future Censuses.

2. Background

The 2010 DRIS Questionnaire Data Capture operation was responsible for the capture and conversion of respondent data from paper questionnaires received from Update/Leave,

mailout/mailback, Update/Leave, The Be Counted operation, and paper-based field operations, as well as, questionnaire inputs from telephone-based questionnaire operations (U.S. Census Bureau, 2010a). The input to DRIS consisted of paper forms (either as postal receipts or as shipments of enumerator questionnaires from Local Census Offices) (U.S. Census Bureau, 2010a). Paper forms were subject to check-in, sorting, document preparation, scanning, data capture (Optical Character Recognition (OCR), Optical Mark Recognition (OMR), or Key-From-Image (KFI) Edits, Document Analysis, Checkout, Warehousing, Destruction, and Quality Assurance prior to integration of captured data (U.S. Census Bureau, 2010a).

Census 2000

In Census 2000, Lockheed Martin was the provider for the Paper Data Capture Systems Contract and was responsible for the Data Capture System 2000 (DCS2000) (U.S. Census Bureau, 2001a). The bid for the paper data capture center operations was a separate contract that was awarded to Thompson Ramo Wooldridge, Incorporated (TRW) as the Data Capture Services Contract (DCSC) (U.S. Census Bureau, 2001a). The telephone operations for Census 2000 were handled under a separate contract awarded to Electronic Data Systems (EDS) (U.S. Census Bureau, 2001b). The National Processing Center's (NPC) Jeffersonville Telephone Center, a Census Bureau facility, participated as one of the call centers used in the EDS solution for Census 2000 (U.S. Census Bureau, 2001b).

In Census 2000, there were four data capture centers, located in Pomona, CA, Baltimore, Phoenix, and Jeffersonville, IN. The table below indicates the volume of forms processed at each site.

Table 2.1, Census 2000 Form Workloads by Data Capture Center

Census 2000 Data Capture Center	Cumulative Forms processed
NPC	32,423,381
Pomona	43,870,967
Phoenix	45,961,899
Baltimore	39,345,532
TOTAL	161,601,779

Source: U.S. Census Bureau 2003a

Mid-decade tests

During tests in the decade leading up to the 2010 Census, including the 2002 Census Test, 2003 National Census Test (NCT), 2004 Census Test, 2005 NCT, 2006 Census Test and 2007 NCT, data capture was never considered an official test objective (U.S. Census Bureau, 2003b). NPC provided data capture using the Census Bureau's legacy solution at the same time and a request for proposal was created that did not parallel the legacy solution. In the future, all decennial testing should take into account what all the final requirements will be. Subject matter analysts

maximized short form content onto the same sized piece of paper, as the Census 2000 short form (U.S. Census Bureau, 2003b, 2006a, 2006b, and 2008a). However for the DRIS contract, data capture Service Level Agreements (SLAs) for overall output accuracy of OMR and OCR were increased in the 2010 contract from those used for Census 2000 prior to testing to ensure requirements could accommodate the budget (U.S. Census Bureau, 2009a) as shown in Table 2.2 (Muenzer, 2011c).

Table 2.2: Service Level Agreements for Overall Output Accuracy in 2000 and 2010

SLA	DCS2000	DRIS 2010
OMR	99.00%	99.80%
OCR	98.00%	99.00%

Sources: Muenzer 2011c, Lockheed Martin 2011a

Recommended forms design guidelines provided by the DRIS contractor in 2006 (as described in the response to Question 13) were not considered until after the final design was already established during the 2006 Census Test and therefore impacted data capture system development and operations (Lockheed Martin, 2010). During the 2005 NCT, one panel attempted to test some of the lessons learned from the DCS2000, including an instruction telling respondents how to correct mistakes (by drawing a line through incorrect entries and writing the corrected response as close as possible to the field in question) and to print in all uppercase. However, there was no evaluation of the impact of these changes and the results of the 2005 NCT were not published in time to inform the 2006 Census Test (U.S. Census Bureau, 2006d).

During mid-decade testing, Census Bureau managers decided that the Census 2000 form design (in terms of physical paper size) would be the baseline for the 2010 Census. The size of paper ultimately selected for the 2010 Census form (also used for Census 2000 short forms), of 8½" x 25" (when unfolded) was too long to work with any of the existing commercial automated scanning products. DRIS had to work with the selected scanner manufacturer to customize their scanning solution that modified the scanner feeder, transport, output hopper and software to accommodate the form size ultimately chosen for the 2010 Census.

2008 Census Dress Rehearsal

For the 2008 Census Dress Rehearsal, DRIS was only partially engaged in data capture and system integration (U.S. Census Bureau, 2009a). There were no field enumeration activities and a decision was made due to budget constraints not to test automated mail check in and sorting. NPC was used as the only paper data capture site for the 2008 Census Dress Rehearsal (U.S. Census Bureau, 2009a). DRIS did not have an opportunity to test their management of staffing or facilities. Also, mail sorting used the Census 2000 Docutronix 2000 mail sorters integrated

into DRIS and maintained by NPC, rather than DRIS chosen equipment (U.S. Census Bureau, 2009a). This meant that the sorting equipment eventually selected for 2010 by the DRIS contractor were never tested in an operational setting until the Operational Test and Dry Runs for DRIS. In the 2008 Census Dress Rehearsal, DRIS tested only a handful of distinct form types (U.S. Census Bureau, 2006c), of over fifty-six distinct form types processed by DRIS in 2010 (U.S. Census Bureau, 2010a).

2010 Census

The 2010 Census DRIS program integrated several Census 2000 contracts including, telephone, paper operations, and paper systems contracts, under one contract called DRIS. DRIS included the telephone systems and operations (Telephone Questionnaire Assistance and Coverage Follow up) (Lockheed Martin, 2008a), paper data capture systems and operations, as well as data normalization and integration tasks performed by Headquarters processing for Census 2000 (U.S. Census Bureau, 2009a). This was a deliberate strategy for 2010 based on recommendations from Census 2000.

When the DRIS contract was awarded in 2005, the plan for 2010 was that handheld computers were originally supposed to electronically capture enumerator interview data from Nonresponse Followup and other field operations. Because this was the plan for 2010, DRIS bid for one fewer data capture center than what was used for Census 2000, under the assumption that there would be reduced processing capacity since there would be fewer paper enumerator forms than Census 2000 under this plan. The response to Question 10 describes in further detail the far-reaching impacts of the change to use paper enumerator forms for 2010 rather than handheld computers.

A cluster-based design was used for DRIS. The cluster design was first developed for Census 2000. The idea was that an independent, isolated processing segment could be developed, tested, controlled and operated in a lab environment on a small scale prior to production. A cluster is a logical unit of hardware and software. Once the cluster design was fully tested and functional, the ability to duplicate and scale up to various production levels was relatively easy to implement and maintain (Muenzer, 2011a). This cluster architecture was critical in scaling up the system for the additional clusters needed to support the Matching, Reinterview and Coding System (MaRCS) as described in the response to Question 10.

For the 2010 Census, Phoenix and Baltimore each had twelve clusters and NPC had six clusters (Muenzer, 2011a). The system was sized for NPC to have one-half the capacity for check in and capture of paper questionnaires as compared to the Phoenix and Baltimore data capture sites. The intended distribution was for forty percent of the work to go to Phoenix (generally from states west of the Mississippi River and Puerto Rico, as well as foreign language forms). Forty percent of the work was planned to go to Baltimore (generally from states east of the Mississippi

River). Twenty percent of the work was planned to go to NPC (generally from the Great Lakes states and Upper Midwest). Table 5.2 shows the actual distribution of the capture workload as it came in to the three data capture centers. Phoenix received 37.21 percent of the work, while Baltimore received 37.67 percent of the work and NPC received 25.12 percent of the work. All mailout experimental forms were planned for capture at NPC except for the D-1(X13)[1] booklet forms that were planned for capture at Phoenix to avoid creating a different envelope design for this form (Lockheed Martin, 2009a) and Bilingual forms were planned for capture at all three sites (Lockheed Martin, 2009a). Group Quarters forms (including experimental GQ forms) were planned to be processed only at NPC while Experimental Enumerator and regular Enumerator Forms were to be processed at all three sites.

2010 DRIS Paper Data Capture Overview

Mail returns received at the data capture center were checked-in (given credit for responding) generally within 48 hours of receipt. This is necessary to provide timely data to the Mail Response Rates Feedback Program and the Daily Mail participation rates. This check-in is accomplished by running unopened forms through a mail sorter to check-in the forms by reading the barcode on the form through a window in the envelope.

Forms whose barcodes could not be read through the envelope window were manually checked-in.

Enumerator form returns are received in a specially designed box, unfolded flat and are sent straight to the document preparation area. This custom box also doubles as a tray for data capture purposes once the lid is removed. Enumerator forms are checked in using barcodes that are read by the scanner. All enumerator forms must be scanned within 10 days of receipt at the data capture centers to feed data to the Matching Review and Coding System (MaRCS) on a timely basis.

The document preparation area removes mail returns from the envelope, unfolds them, places in trays for scanning. Also at this stage, booklet forms such as the bilingual form and the D-1(X13) booklet are sent to the guillotine to remove the binding. Boxes of enumerator returns are examined, and are prepared for scanning. Any damaged forms are transcribed onto a new blank form of the same type and a new barcode label that retains the original barcode is printed on these transcribed forms at this stage of the process.

[1] During the Census, there were several experiments with unique forms. Except for the Bilingual form and this D-1(X13) Booklet form, most mailback forms unfolded to a single sheet of paper. The D-1(X13) was intended to test a different coverage treatment. A description of each form type and form number can be found in Appendix B: Form Type Descriptions By Form Number.

From this point, materials are sent to be scanned. Once scanned, the physical paper forms are sent to warehousing, awaiting confirmation that the data have been received, which authorizes destruction.

For further details about subsequent steps in the data capture process as well as additional detail about this step, consult the DRIS Operations plan (Lockheed Martin, 2008d).

3. Methodology

Table 3.1 below describes the methodology and source data that will be used to answer each research question (as numbered in Section 3.1). An "X" in the column indicates the data source or methodology will be used to answer each question.

Table 3.1 Data Sources/Methodology for Each Research Question

Research Question #	Examination of Cost & Progress Data	Qualitative Analysis of what occurred	Examination of Lessons Learned Documentation	Examination of Contractor data
1	X	X		
2	X	X		X
3		X	X	X
4		X	X	
5	X	X		
6	X	X		
7		X	X	X
8		X	X	X
9		X	X	X
10		X	X	X (including Contract CRs)
11		X	X	X (including Contract CRs)
12		X	X	X
13		X	X	X
14		X	X	X
15		X	X	X

Questions

The DRIS Paper Questionnaire data capture assessment will address the following questions:

3.1. Question 1: What was the cost per paper form processed by DRIS at each site and what was the total average cost per form?

3.2. Question 2: How did the workload estimates in the contract compare with the final actual workloads for each form type processed by DRIS?

3.3. Question 3: How was staffing impacted by differences between the actual receipt curve and the assumed receipt curve provided in the contract?

3.4. Question 4: What changes necessitated a change in the original workload estimates?

3.5. Question 5: What was the reverse check-in rate?

3.6. Question 6: What was the manual/exception check-in rate?

3.7. Question 7: How did reprocessing of ambiguous responses affect systems and operations?

3.8. Question 8: How did field procedures and materials impact data capture?

3.9. Question 9: Were there any requirements which were not clearly defined so that the contractor could provide the intended outcome?

3.10. Question 10: What was the impact to Paper Data Capture of late and or ad hoc requirements presented by the Census Bureau to the DRIS contract?

3.11. Question 11: How did the addition of late requirements covered within the existing contract scope impact budget and resource scheduling?

3.12. Question 12: Were the Census Bureau deliverables to DRIS timely and complete? If not, what was the effect?

3.13. Question 13: Were the deliverables from the contractor to the Census Bureau timely and complete?

3.14. Question 14: What were the impacts of changes on interface management—what worked and what did not? We will try to quantify this where possible.

3.15. Question 15: Were the testing methodologies used to test systems and operations sufficient to meet requirements?

4. Limitations

Because DRIS is an integrated contract, that includes telephone operations, paper operations, data integration operations, equipment management, facilities, support and a variety of other costs, there are shared costs across the DRIS enterprise that cannot be solely attributed to paper questionnaire data capture (Pentercs, 2011). Some cost calculations account for shared costs across the various DRIS components (Pentercs, 2011).

The Decennial Statistical Studies Division (DSSD) has not committed resources to this assessment, so only Cost and Progress data and contractor reports/materials can be examined.

This assessment will not examine the Decennial Response File, the Census Unedited File, or Universe Control and Management data. This assessment can only examine aggregate level counts.

5. Results

5.1. Question 1: What was the cost per paper form processed by DRIS at each site and what was the total average cost per form?

The cost results presented in this assessment were generated by program office staff using methods predating the US Census Bureau's commitment to comply with Government Accounting Office's cost estimating guidelines and the Society of Cost Estimating and Analysis best practices. Hence, while the Census Bureau believes these cost results are accurate and will meet the needs for which they will be used, the methods used for estimating costs of 2010 Census operations may not meet all of these guidelines and best practices. The Census Bureau will adhere to these guidelines in producing 2020 Census cost estimates.

DRIS was an integrated contract that shared components with other data collection operations such as telephony. These costs cannot be considered as stand-alone costs and cannot be generalized to any future operation (Pentercs, 2011). There were cost savings utilized by integrating these segments into one contract (Pentercs, 2011).

Direct Labor Cost for form face image processing per case

To determine a reasonable cost per case, calculations for paper data capture costs are based on the number of form face images processed by DRIS, because different form types have different numbers of images. Most forms have two images each; booklet forms have twelve images each. Different form types have different data densities and may have contained a different amount of data to be keyed so, in addition, some forms may have had no images presented to a keyer and would require less cost to process.

Based on Cost and Progress information, the direct labor costs, to process forms during 2010 (Service Contracting Act labor for Phoenix/Baltimore, government payroll for NPC) are displayed in Table 5.1.

Table 5.1 Average Direct Labor Cost* Per Form Face Image Processed By DRIS

Data Capture Center	Total Form Face images processed in Calendar Year 2010	Total Direct Labor Cost*	Total Hours Worked	Effective Hourly Pay Rate	Average Cost per form face image
NPC	88,958,008	$13,792,462	960,690	$14.36	16¢
Baltimore	150,550,862	$42,132,617	1,633,062	$25.80	28¢
Phoenix	162,154,980	$58,609,575	1,892,392	$30.97	36¢
TOTAL	**401,663,850**	**$114,534,654**	**4,486,144**	**$23.71**	**29¢**

Source: Calculated from Cost and Progress Data

*NOTE: Direct Labor costs include the cost of front line staffing at each site for all costs incurred during the production time window. The effective rate is a blended rate of part time and full time employees and straight time and overtime pay calculated by dividing the total cost by the total number of hours worked at each site.

Using this approach, the average direct labor cost per form face image processed in calendar year 2010 was 29¢ across all three data capture sites. However, the average direct labor cost per image processed by DRIS sites (Baltimore and Phoenix) alone was 32¢ overall. The total costs yield an average effective pay rate of $23.71 per hour across all three data capture centers. This excludes all system, hardware, or integration costs and is only a rough approximation of the amount of direct labor it took to process each form image, assuming all images require an equal amount of direct labor cost. These costs represent only the hours of front line keyers, clerks, or equipment operators and exclude management or overhead costs.

NPC costs were less primarily due to the difference in labor costs. If one were to examine the total number of images processed divided by the total number of hours worked, the numbers are very close between NPC and DRIS sites. If one were to apply the same effective hourly pay rate to NPC that DRIS paid, the direct labor costs per image are very comparable between DRIS and NPC if you control for the two weeks of down time at Phoenix and Baltimore (as described in the response to Question 3). Because DRIS was mandated to pay the Service Contracting Act labor rates in Phoenix and Baltimore, labor costs were significantly higher for forms processed at those sites because the hourly labor rates were higher.

NPC paid much lower labor rates than the Service Contracting Act rates that the contract was required to pay at the Phoenix and Baltimore sites. The biggest share of the difference in cost for NPC appears to be the difference in labor rates. When controlling for the difference in labor rates between NPC and Baltimore and the downtime at other sites, the direct labor costs are very comparable between NPC and Baltimore sites.

There was a period at the beginning of field enumeration where the Paper Based Operations Control System (PBOCS) field shipping system had a significant shipping backlog (as described in the response to Question 3) and the paper data capture centers were not receiving questionnaires to process in the volume expected. NPC was able to reassign staff to other 2010 Census projects (such as kit preparation and geographic keying).

The Phoenix and Baltimore paper data capture centers did not have the flexibility of reassigning staff to other 2010 Census projects. However, they did try to change schedules and offer liberal leave to staff to avoid paying them during slow days. See Question 3 for a description of the additional costs borne by Phoenix and Baltimore during this period.

The distribution of form types processed by each site varied, which could impact the cost differences across sites. Table 5.2 shows the distribution of the three largest volume form types for each site. NPC was intended as the only site to process GQE forms; however, a small number of these forms were erroneously directed to other sites. NPC processed a smaller number of Bilingual forms than other sites because the geographic area assigned to NPC for data capture contained fewer households selected for the Bilingual mailing than for the geographic areas covered by Phoenix or Baltimore sites. Table 5.2 shows the distribution of the overall largest volume form types that was processed at each data capture center. The third most processed form at NPC was not the Bilingual (Booklet size) questionnaire, as it was at other sites. The third largest volume form processed at NPC the Individual Census Report.
Table 5.3 shows the distribution of the volume of physical paper size of forms processed at each site.

Table 5.2: Distribution of Overall Number of Largest Volume Form Types Processed by Site March 4, 2010 through September 7, 2010

Site	# D-1[2] Forms Processed	% of Total D-1 Forms Processed	# D-1(E/S)[3] Forms Processed	% of Total D-1(E/S) Forms Processed	# D-1(E)[4] Forms Processed	% of Total D-1(E) Forms Processed
Phoenix	27,611,787	33.30%	4,494,248	61.79%	21,128,673	36.87%
Baltimore	34,994,541	42.20%	2,297,945	31.59%	24,492,729	42.74%
NPC	20,323,250	24.51%	481,564*	6.62%	11,690,181	20.40%
Total	**82,929,578**	**100.00%**	**7,273,757**	**100.00%**	**57,311,583**	**100.00%**

Source: Cost and Progress Data

*The D-1(E/S) was not one of the largest volume forms at NPC, since the geographic area captured by NPC had fewer of these forms than Baltimore or Phoenix.

Table 5.3 shows the proportion of the total workload at each site by the physical paper size of the forms processed. The table shows that NPC processed 99.99 percent of all 8½" x 11" (GQE) forms, while Phoenix processed the largest share of booklet forms (Bilingual and D-1(X13) forms) at 61.90 percent. The maximum scanning rates tested were one hundred twenty-eight forms per minute of the 25 ½ "x 11" size and three hundred twenty-four forms per minute of the 8 ½" x 11" size. Booklet forms such as the D-1(X13) and D-1(E/S) have an additional labor overhead of having the spines removed at the guillotine.

[2] The D-1 form is the standard Mailout/Mailback form also used for Update/Leave areas. A description of each form type and form number can be found in Appendix B: Form Type Descriptions By Form Number.

[3] The D-1(E/S) form is the bilingual form.

[4] The D-1(E) form is the standard Enumerator form.

Table 5.3 Form Size Proportion of Total Workload by Site March 4, 2010 through September 7, 2010

Form Size	Phoenix (number)	Phoenix (% of total)	Baltimore (number)	Baltimore (% of total)	NPC (number)	NPC (% of total)	TOTAL (number)
8½" x 11"	5	0.00%	712	0.01%	8,051,060	99.99%	8,051,777
11" x 25½"	58,499,640	37.99%	61,485,477	33.93%	34,012,904	22.09%	153,998,021
Booklet	4,515,569	61.90%	2,298,207	31.50%	481,627	6.60%	7,295,403
Unknown	575	10.02%	4,389	76.46%	776	13.52%	5,740
Total	63,015,789	37.21%	63,788,785	37.67%	45,546,367	25.12%	169,350,941

Source: Calculated from Cost and Progress Data

The costs reported in Table 5.1 also exclude Group Quarters Validation (GQV) processing and the 2008 Census Dress Rehearsal. Since only NPC was involved in processing GQV and the 2008 Census Dress Rehearsal, such a comparison does not apply to those operations. DRIS labor rates for the contracted paper data capture centers were determined by the U.S. Department of Labor, and Phoenix had a higher hourly labor rate, followed by Baltimore. NPC had government employees and their pay rates were established by the U.S. Office of Personnel Management. DRIS was required to pay wages based on a U.S. Department of Labor wage rate determination for Service Contract Act labor.

The effective pay rate displayed in Table 5.1 was calculated from Cost and Progress (C&P) report data using total direct labor cost divided by total hours and represents a blended rate of part time and full time employees and straight time and overtime pay across all labor categories. This rate only includes direct labor and does not include indirect costs or overhead costs

The total Estimate at Completion cost (as of July 2011) for the paper data capture was $473,642,000 (Pentercs, 2011). The Cost and Progress data show that the paper channel processed 166,415,630 total paper forms (including both GQV and 2010 production). C&P data show that NPC reported $15,928,660 in total direct labor cost for both GQV and 2010 Production consisting of $2,231,516 for GQV and $13,792,462 for 2010 production.

If the total costs for DRIS operations and staffing (Contractor costs plus NPC labor costs) as itemized in Table 5.4 are divided by the total number of paper forms processed by DRIS as shown in Table 5.5, this yields an average data capture cost of approximately $2.94 per paper form processed through data capture in DRIS.

Table 5.4 Total Data Capture Cost

NPC GQV Direct Labor Cost	$2,231,516
NPC 2010 DRIS Direct Labor Cost	$13,792,462
Total Estimated DRIS Paper Channel Cost*	$473,642,000
Total Estimated Data Capture Cost	**$489,665,978**

Source: Pentercs 2011, Cost and Progress Data

*NOTE: The DRIS Paper Channel Cost includes systems, overheads, hardware, management, and other integrated costs that cannot be separated from other parts of the contract.

Table 5.5 shows the total number of paper forms processed using the DRIS system for the 2008 Census Dress Rehearsal, GQV, and the 2010 Census. Appendix A shows a detailed distribution of the GQV forms and 2010 forms. The table in Appendix B describes the form types listed in Appendix A by form number. For Dress Rehearsal, DRIS did not provide Cost and Progress information, so C&P had to rely on data provided by Headquarters Processing for the 2008 Census Dress Rehearsal. Table A.3 in Appendix A shows the distribution of Dress Rehearsal forms provided to C&P by Headquarters Processing.

Table 5.5 Total Number of Paper Forms Processed by DRIS/NPC in Production

Phase	Total paper forms processed by DRIS/NPC
2008 Census Dress Rehearsal (2008)	255,874
Group Quarters Validation (2009)	2,067,511
2010 Census (through September 7, 2010)	164,353,859
Total	**166,671,504**

Source: Cost and Progress Data

Total DRIS Contract Cost

Because DRIS is integrated across several program areas, an overview of all DRIS costs is presented below. Only Paper and Workflow Control and Management (WCM) segments of DRIS are within the scope of this assessment. Separate assessments will address TQA and CFU contract cost.

The following represent total contract costs calculated as of March 31, 2011 (Pentercs, 2011):

- Total Contract Value: $1,019,134,529 (initial contract award was October 2005)
- Estimate At Completion (EAC): $980,203,000 (As of the July 19, 2011 DRIS Contract Performance Report with data through June 24, 2011)

5.2. Question 2: How did the workload estimates in the contract compare with the final actual workloads for each form type processed by DRIS?

DRIS used paper data capture questionnaire workload estimates for system sizing. These estimates were developed using a "worse case" scenario, and as a result the estimates overstated the final actual workloads encountered for most questionnaire form types. When the replacement mailing strategy was changed to a targeted and blanket approach with fewer total replacement questionnaires to be sent out, the capture workload estimates for replacement questionnaires were never revised accordingly. Cost and Progress data through September 7, 2010 show that in total, DRIS captured about 87.64 percent of the estimated workloads for DRIS operations conducted in Calendar Year 2010 (See the tables in Appendix A) (U.S. Census Bureau, 2011a).

In general, a ten percent design capacity contingency was initially figured into the original data capture workload estimates. Since DRIS received about 12.36 percent less than the design capacity prescribed in Contract Section J.27, the workload estimates were off by about 2.36

percent from the true final workloads encountered (if the ten percent design capacity contingency included in contract Section J.27 is discounted).

This could be partly explained by the fact that in planning workload estimates, the number of households that would send back both an initial and replacement questionnaire was overestimated. Table 5.6 shows a high level comparison of the estimated workloads of forms compared with the final workloads encountered. The tables in Appendix A show detailed distributions for each form type.

- For GQV operations in 2009, DRIS captured 2,067,511 forms at NPC (U.S. Census Bureau, 2011a) against an original workload estimate of 2,070,000, or about 99.88 percent of the originally estimated workload (U.S. Census Bureau, 2011a). It should be noted that while the total estimated number of GQV forms captured was close to the projected workload estimate, the distribution of the single sheet as compared with booklet forms actually processed was not as close to the original estimate, as seen by Table A.2 in Appendix A.

- For 2010 Census paper production operations, DRIS captured a total of 164,348,119 forms (excluding forms identified by DRIS as "unknown") through September 7, 2010 (U.S. Census Bureau, 2011b) against an original workload estimate of 187,524,117 forms (U.S. Census Bureau, 2011b).

- For 2010 Census paper production operations, DRIS processed over three billion individual checkbox fields on Census Questionnaires using Optical/ Mark Recognition software.

Table 5.6 Total Paper Forms Data Captured in DRIS 2009-2010

Total 2010 Census Forms Data Captured	Estimated Data Capture Workload	Actual Data Capture Workload	Percent of Estimate	Notes
Total 2009 Forms Data Captured	2,070,000	2,067,413	99.88%	Although the actual is close to the estimate, the distribution of individual form types is not as close.
Total 2010 Forms Data Captured*	187,524,117	164,348,119	87.64%	Estimates were never revised in light of the change to the replacement mail strategy.
TOTAL Forms Data Captured 2009-2010	189,594,117	166,415,532	87.78%	The estimates in section J.27 of the contract deliberately overstated the expected workload by 10% as a contingency.

*Excludes form types classified as "unknown" by DRIS
Note: GQV forms were captured in 2009, 2010 Census Questionnaires were captured in 2010.
Source: Calculated from Cumulative Data Capture Cost and Progress report and Section J.27 of the DRIS contract

These numbers are data capture check-in counts representing actual pieces of paper processed by DRIS through September 7, 2010. This means that the same housing unit could have multiple forms, and each form is counted in this number. For GQV operations, one Other Living Quarters (OLQ) could have multiple forms also (such as a continuation form or housing unit form). DRIS was required to process every questionnaire received, regardless of whether it was a duplicate of another form from the same housing unit or OLQ.

When examining the data by form type, the form with the highest percentage of the expected workload was the D-1(XB), an experimental form that was a control panel designed to test over-count coverage questions. About 179.54 percent of the expected workload in Contract Section J.27 was received for this form. This is because the experimental design was revised with a larger sample size after workloads had already been placed in Contract Section J.27 and the workload estimates were never increased accordingly for this form type.

The 2010 Census was the first time that Enumeration at Transitory Locations was conducted as a separate stand-alone operation. The forms for this operation (D-15 and the D-15PR(S)) were processed at levels well below the projected capture workloads because of the difficulty of developing reliable workload estimates for this operation.

The tables in Appendix A – Comparison of Estimated and Actual Data Capture workloads show the workload comparisons for each form type for 2010 Census processing as well as GQV processing. The table in Appendix B describes the form types listed in Appendix A by form number.

5.3. Question 3: How was staffing impacted by differences between the actual receipt curve and the assumed receipt curve provided in the contract?

DRIS was originally set up to interface with PBOCS to receive FedEx shipment information that detailed the Box and Census Form IDs they were expected to receive (Lockheed Martin, 2011a). After a configurable timeout period, DRIS would investigate any missing materials (Lockheed Martin, 2011a). Because of PBOCS performance problems, the Census Integration Group (CIG) made a decision to abandon PBOCS for shipping and use a contingency system developed by the Administrative and Management Systems Division (AMSD) for shipping materials.

Local Census Offices were unable to ship materials for the first several days of NRFU, and materials were accumulating in the LCOs and not being shipped to DRIS on schedule. Census Bureau senior leadership made a conscious decision to remove shipping functionality from PBOCS, which negatively affected the ability of DRIS to staff properly for the expected workload.

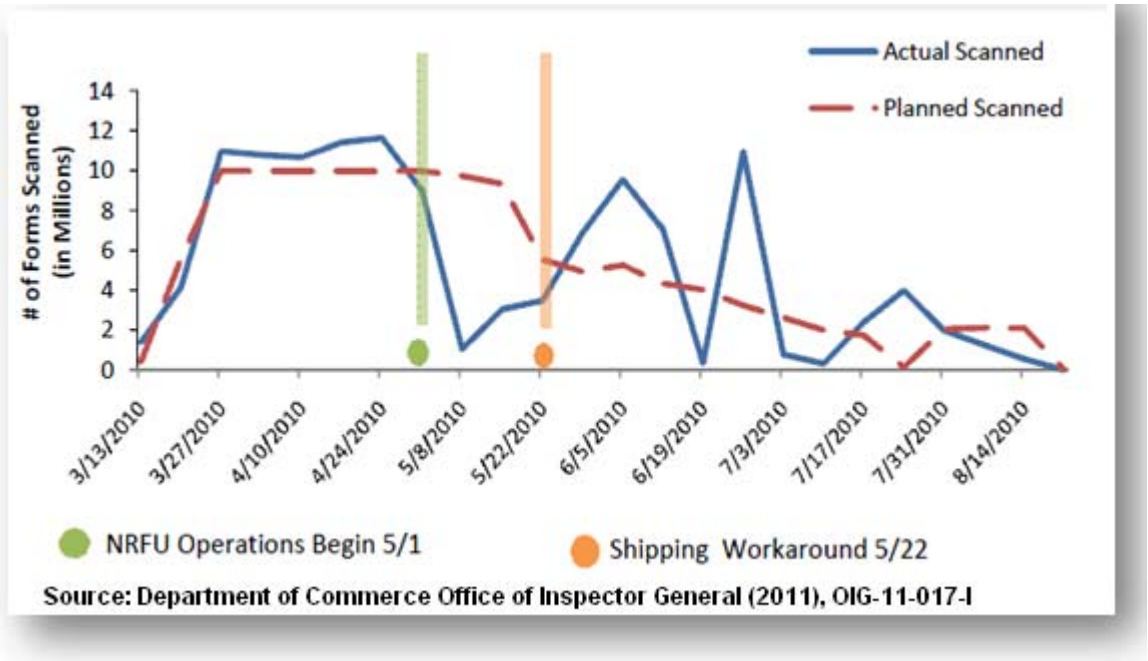
Since boxes of Nonresponse Followup (NRFU) forms were not received on schedule at the Data Capture Centers, the DRIS Senior Leadership Action Panel (SLAP) decided on May 26, 2010 to continue staffing at full levels while waiting for PBOCS shipping backlogs to be resolved. According to the SLAP Operational Contingency Log, this cost was approximately $700,000 per week at each of the two DRIS-managed data capture sites, or approximately $1,400,000 total for Phoenix and approximately $1,400,000 total for Baltimore for a total estimate of $2,800,000 (Lockheed Martin, 2010b). Additional NPC staffing costs for this period could not be quantified because NPC was able to reassign staff to other projects (such as kit preparation and geographic keying) during this period.

Since the AMSD shipping contingency did not have the shipping information in the interface with DRIS, much of the effort DRIS spent to develop this reconciliation feature was lost and significant additional money had to be spent on performing extra forms reconciliation at the end of operations as described in the response to Question 14 (Lockheed Martin, 2011a). Also, since the shipping contingency did not communicate what would arrive in advance, DRIS attempted to coordinate with FedEx to get an idea of what volume of materials to expect for the next day's processing. The original plan would have allowed DRIS a one-day advance window to set staffing levels with enough advance notice to know when to expect materials. Because the shipping contingency did not give DRIS advance notice of what workload to expect on any given day, this required an increase in both the communication effort and level of coordination needed

between DRIS and FedEx. There was already a modest budget in DRIS to support end of program reconciliation tasks, but this had to be increased because the decision to remove shipping functionality from PBOCS resulted in forms and boxes which were unaccounted for that could have otherwise been tracked using PBOCS had that functionality worked as designed.

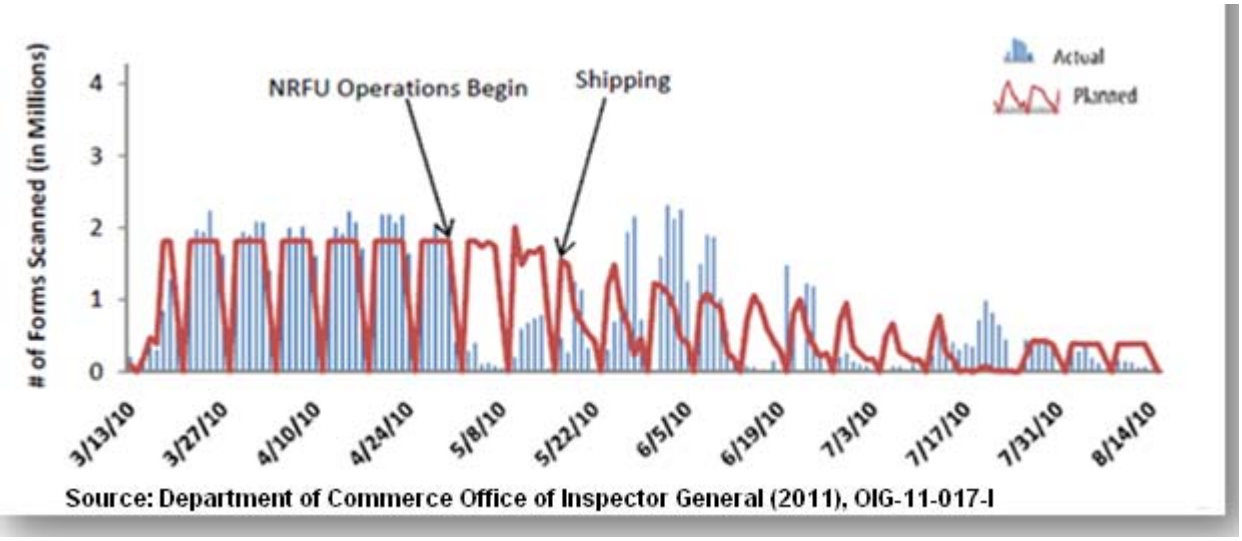
The Department of Commerce Office of Inspector General (OIG) issued a report (2011) that included the graphs in Figure 1 and Figure 2. The graph in Figure 1 shows how the actual volume of forms scanned fell off dramatically once NRFU operations began in the field. In this graph, the dashed red line represents the assumed receipt curve provided in the contract (at the weekly drill-down level) and the solid blue line represents the actual receipt curve. DRIS was instructed to staff to the plan during this time. You can see from Figure 1 the volume of materials that did not come into DRIS for scanning during the time period between the start of NRFU operations on May 1, 2010, and the time when the shipping work around was deployed on May 22, 2010. This delay added to the risk of the timely completion of the reprocessing effort described in the response to Question 7.

Figure 1 Actual Forms Scanned and Planned Forms Scanned (weekly)



The graph in Figure 2 shows a daily overview of the entire data capture production window. The white space under the plan line is again clearly visible for the time period that PBOCS was unable to handle the volume of materials requiring shipping. As noted above, the SLAP directed that full staffing levels be maintained during this period. These staffing levels resulted in the lost opportunity cost of being able to capture expected volumes which did not materialize until later in the data capture process.

Figure 2: Actual Forms Scanned and Planned Forms Scanned (daily)



The U.S. Department of Commerce Office of Inspector General (2011) stated the following regarding the shipping situation:

> *We did not identify any significant issues with the [Paper Data Capture Centers' (PDCCs')] handling of the unexpected volumes and erratic delivery of forms to be scanned during NRFU. The [Local Census Offices] had significant difficulties shipping questionnaires in a timely manner to the PDCCs during NRFU due to breakdowns in PBOCS; this resulted in days of unexpected downtime at the PDCC from mid-April onward. On May 22, the shipping functionality was moved from PBOCs to another management system [provided by AMSD] because of the negative impact on shipping. Under [Census] Bureau direction, DRIS contractors revised the overall staffing plans to compensate for some unproductive hours. Even with these adjustments, the contract stayed within budget during this period and continued to meet key performance indicators. Had PBOCS problems not slowed down the delivery of questionnaires to the PDCC, it is conceivable that labor costs may have been lower during these months.*

5.4. Question 4: What changes necessitated a change in the original workload estimates?

The design for the various experimental paper forms was unknown when the workloads were first created for Contract Section J.27. Several change requests (CRs) were issued to increase workloads for experimental forms. Because the low volume did not justify the cost, the Group Quarters Enumeration Operational Integration Team issued a CR to use the Stateside Military Census Report (D-21) for Puerto Rico and also to use the stateside Shipboard Census Report (D-23) for Puerto Rico (Ciango, 2009). This meant that the original forms workloads projected for Puerto Rico versions of these forms were combined into the stateside workloads for the D-21 and D-23 when data were displayed in final reports. The Experimental Individual Census Report, D-20(X1), was not part of the original plan, and in actual production reduced the number of D-20 forms processed. See the notes section of the table in Appendix A for more information.

There appeared to be an error in the Contract Section J.27 workload estimates for D-1PR(S). The original workload estimate for this form type was somehow lost when the form naming conventions changed. The tables in Appendix A show the workload comparisons for each form type for 2010 Census processing, as well as GQV processing. The table in Appendix B shows a description of each form type.

For the standard stateside D-1 mailout/mailback forms, the original workload estimates were never revised downward in light of the change to the replacement mailing strategy. The original plan was to send out about forty million replacement questionnaires. As a result of the change in the replacement mailing strategy, a total of 10,000,000 targeted replacement questionnaires and 24,764,056 blanket replacement questionnaires were sent out, resulting in a total of 34,764,056 replacement questionnaires (Smith, 2010), well below the original estimate of forty million.

5.5. Question 5: What was the reverse check-in rate?

All mail returns to the data capture centers were required to be "checked in" within 48 hours of receipt. This was accomplished generally by using high speed sorters, without removing forms from the envelope. The sorters checked the forms in by reading the Census ID barcode through the envelope window. The mail returns were "checked in," and the housing unit was given credit for responding and thus the housing unit was then removed from the Nonresponse Followup (NRFU) workload. The check-in information was also used for the response rates reports. Once forms were opened and data captured, they were "reverse checked in" if they were discovered to be "blank" questionnaires or had insufficient data as defined by the Census Bureau (Lockheed Martin, 2011a) to be included in the census. For these questionnaires, the housing unit check-in database was then updated to show that the housing unit did not respond to the census and that it was still eligible for field enumeration.

Through September 7, 2010, DRIS processed 1,286,167 reverse check-in records (Lockheed Martin, 2011a). Of a total of 92,762,827 mail returns eligible for reverse check-in (RCI) (U.S. Census Bureau, 2011a), this represents a reverse check-in rate of approximately 1.39 percent.

Tests earlier in the decade saw these rates in a range of between one and two percent (U.S. Census Bureau, 2005). The way reverse check-ins were processed continuously evolved such that direct comparisons are not possible with Census 2000 or mid-decade tests. The forms used in this equation only include Master Address File (MAF) based IDs, since only forms with MAF based IDs are eligible for reverse check-in (U.S. Census Bureau, 2005). For a detailed list of form types eligible for reverse check-in, see the Blank Form Specification section in the Data Quality Management Plan and Production Results (Lockheed Martin, 2011a). DSSD will give a more detailed analysis of reverse check-in data in the Universe Control and Management/Response Processing Systems Assessment based on household level data.

5.6. Question 6: What was the manual/exception check-in rate?

Of all mail return form types eligible for sorter check-in (as enumerated in Appendix D), the sorters were able to check-in a slightly lower percentage than originally planned. Cost and Progress data show that about 3.11 percent were manually checked in (U.S. Census Bureau, 2011b) compared with a worse case modeling assumption of 3.9 percent (Lockheed Martin, 2009b).

Generally, mail return forms are checked in at the high speed sorters if the Census ID barcode shows through the window of the envelope. If the Census ID barcode cannot be read at the sorter, these forms are sent to a reject pocket in the sorter. Rejected forms are inspected then run through the sorter a couple more times. The remaining rejects are then taken to manual/exception check in to be processed. If forms are received in nonstandard respondent supplied envelopes and rejected, these are also sent to manual/exception check-in. Also, on days with a comparatively low volume of mail, forms may be taken directly to manual/exception check-in (MCI).

The table in Appendix D (Exception Check-In Rates for Mail Returns by Form Type) lists the exception check-in rate for all eligible form types. Eligible form types are those that are mail return form types eligible for sorter check in. Enumerator returns or field supplied returns are not checked in until they are scanned at document imaging and do not go through the sorters. The table in Appendix B shows a description of each form type.

Due to the higher than expected exception check-in rates during the 2008 Census Dress Rehearsal, the return envelopes for the standard D-1 short form were re-designed for the 2010 Census. The redesign allowed for a form to be inserted upside down into the return envelope, as long as the barcode was on the correct side (facing the window). These forms could be read by

the sorter (DeMaio, Beck, Schwede, 2008) resulting in fewer exception check-ins. Some form types experienced much higher exception check in rates than other forms as seen in the table in Appendix D. These forms did not have the redesigned envelope. Only the standard stateside D-1 mailout/mailback form had the redesigned envelope.

Also, at the recommendation of the Statistical Research Division (SRD), an illustration was added to the back of the envelope showing respondents how to correctly insert the form in the envelope, so that barcodes could be seen through the envelope window (DeMaio, Beck, Schwede, 2008). The D-1 forms had the lowest exception check-in rate of form types eligible for sorter check-in. The D-1 forms are the only forms which had the envelope redesign. Forms without the envelope redesign experienced higher exception check-in rates, so the data would seem to indicate that the envelope redesign had a positive impact towards reducing the exception check-in rate for forms that used the redesigned return envelope.

There were a large number of Puerto Rico mail returns requiring manual check-in (U.S. Census Bureau, 2011c). The DRIS Program Management Office believes that this may be due in part because the graphic that SRD recommended be added to the envelope was not translated into Spanish, and appeared in English on the Puerto Rico envelopes even though the questionnaire itself was written in Spanish (U.S. Census Bureau, 2011c). Puerto Rico forms used the Update/Leave methodology. As such, they were not mailed out and they did not have postal tracking. Since they did not have the need for a barcode that could be used for postal tracking, the return envelope for these forms did not have the double window redesign, meaning that there was only one way that the form could correctly be inserted in the return envelope to be read by the sorter. It is recommended for the future that messaging on the envelopes should be in the same language as the questionnaire itself.

Another observation from the DRIS Program Management Office is that that many Puerto Rico respondents simply did not understand the tenure (home ownership) question. As such, many Puerto Rico respondents sent forms via certified mail in non-standard envelopes, with mortgage deeds, etc., attempting to prove home ownership (U.S. Census Bureau, 2011c). More research or cognitive testing should be conducted with the translation of that question to avoid these kinds of issues in the future (U.S. Census Bureau, 2011c).

DRIS ran a separate tally for the Data Quality Management Plan and Production Results and came up with a somewhat different value for the manual check-in rate. If a form was so physically damaged that it could not be run through the sorter or scanners, it was transcribed onto a new form and manually checked in. The counts DRIS supplied to Cost and Progress for exception check-in included these transcribed forms. DRIS provided two separate rates, both including and excluding these transcribed forms.

MCI Rate (with transcription) =

$$\frac{Total\ Number\ of\ Manual\ CheckIn\ Forms}{Total\ Number\ Sorter\ CheckIn\ Forms + Total\ Number\ MCI\ Forms + Total\ Number\ CheckIn\ Scanner\ Forms - Total\ Number\ UAA\ Forms}$$

MCI Rate (without transcription) =

$$\frac{Total\ Number\ Manual\ CheckIn\ Forms - Total\ Number\ Transcribed\ Forms}{Total\ Number\ Sorter\ CheckIn\ Forms + Total\ Number\ MCI\ Forms + Total\ Number\ CheckIn\ Scanner\ Forms - Total\ Number\ UAA\ Forms}$$

MCI Rate with transcription=	3.13%
MCI Rate without transcription=	2.93%

The Data Quality Management Plan and Production Results also produced statistics on the manual check-in rate for Undeliverable-As-Addressed (UAA) questionnaires. Normally, these are checked in at the sorter, but sometimes the United States Postal Service (USPS) placed stickers over the Census ID barcode, or lined through this barcode such that the barcode could not be read by the sorters. Cost and Progress data show that there were a total of 463,919 UAA forms requiring manual check-in (U.S. Census Bureau, 2011d). Of a total of 19,594,395 UAAs checked in (U.S. Census Bureau, 2011d), this represents a UAA MCI rate of approximately 2.37 percent. Generally, these questionnaires are returned in the same outgoing envelope and have not been touched by a respondent (U.S. Census Bureau, 2011c).

Note: The table in Appendix C shows UAA exception check-in rates by form type. The table in Appendix B shows form descriptions by form type.

5.7. Question 7: How did reprocessing of ambiguous responses affect systems and operations?

DRIS implemented several measures to ensure the quality of census results. In particular, one system capability involved the reprocessing and retransmission of data from census paper forms. This capability, which included the generation of a "new" batch of images previously processed then run through each stage of the capture process after imaging (including automated capture by OCR and OMR, keying, and audit resolution and edits) proved vital to ensure that the correct information was processed in the census (Cardella, 2010).

Data Quality was a major focus of the DRIS program (Lockheed Martin, 2011a). An independent contract was awarded to Gunnison and its teaming partner, Advanced Document Imaging, LLC (ADI) to provide Independent Verification and Validation services for paper data

capture. ADI developed the Paper Production Data Quality (PDQ) system and provided expert data analysis, which was essential to ensure that an "independent" check of DRIS data quality was performed. The PDQ process assisted by using a variety of both Commercial Off The Shelf tools, custom software and manual efforts to help identify clustering of errors that DRIS could then investigate and resolve.

Through observation and analysis of 2010 Census data capture, DRIS, along with the help of PDQ, identified scenarios which justified reprocessing for certain fields on the questionnaire, such as population count, race, and Hispanic Origin. In general, each scenario outlined a condition that resulted in clusters of potential errors in the data that had been previously captured and transmitted to headquarters. While each of these scenarios represented a particular deficiency in the accuracy or interpretation of the data captured to a particular respondent relationship, collectively they did not threaten the overall accuracy requirement defined in the DRIS contract. Nevertheless, DRIS management and the Census Bureau felt these systematic errors were serious enough to warrant a manageable effort to process the forms again after the appropriate procedural and system fixes were implemented (Lockheed Martin, 2011a).

During production, several types of reprocessing were performed for various reasons. A total of 3,366,130 forms were reprocessed (Lockheed Martin, 2010d).

While the reprocessing events are documented more thoroughly in the Data Quality Research Agenda, the discussion below outlines the reprocessing efforts:

- D-1(XA)[5] Reprocessing: During production, a problem was found with the middle initial fields on the D-1(XA) forms, namely that the physical location of the middle initial write-in box as printed on the form was not the same as the form defined template. The paper team instituted a fix to match the form definition to the paper form, but due to the sensitivity of the form being an experimental form, DRIS decided to reprocess all D-1(XA) forms regardless of vulnerability to the anomaly (Lockheed Martin, 2010d).

  Reprocessing of the D-1(XA) forms yielded a significant increase in the OCR Accuracy of the Middle Initial field. It increased from 96.40 percent to 98.92 percent because of the reprocessing. In addition, due to a more mature system at the time of the reprocessing, an increase in overall keying and OCR accuracy was achieved (Lockheed Martin, 2010d). A total of 159,651 D-1(XA) forms were reprocessed for middle initial accuracy (Lockheed Martin, 2010d).

[5] The D-1(XA) form was an Experimental form used for Control Panel 1. This form had over count coverage questions. A description of each form type and form number can be found in Appendix B: Form Type Descriptions By Form Number.

- D-1PR(S)[6] Reprocessing: During production, a problem was found with the middle initial fields on D-1PR(S) forms, namely, disproportionate error was found when compared to other field types of this form. For example, many respondents wrote in all three initials of their full name rather than their middle initial only (Lockheed Martin, 2010d). As a result, keyers would key all three, while OCR would select one. The system was updated to select the middle initial of the data captured (Lockheed Martin, 2010d).

  The behavior of Spanish speaking respondents entering all three initials was identified during cognitive testing of a bilingual form in 2007 (Goerman et al., 2007), but the recommendations they provided to address this were not tested or implemented. A similar pattern was observed during the 2010 Census on stateside D-1(E/S) bilingual forms in the Spanish swim lane. However, the data on these forms were too sparse to identify a systematic clustering of error that warranted reprocessing of these forms.

  The main reason that it was not possible to research and implement the recommendations that resulted from cognitive testing is because there was not sufficient physical space on the form to accommodate all of the text required to convey the recommended instruction.

  All affected D-1PR(S) forms processed prior to the fix being implemented were reprocessed (Lockheed Martin, 2010d). The Data Quality team sampled the reprocessed forms and verified the objective was accomplished. In addition, due to a more mature system at the time of the reprocessing, an increase in overall keying, OCR, and middle initial field OCR accuracy was achieved for this form (Lockheed Martin, 2010d). A total of 443,666 D-1PR(S) forms were reprocessed for middle initial accuracy.

- Marginalia Reprocessing: During production, there were significant changes to the defined "rules" that were applied by the edits analysts with respect to the handling of respondent or enumerator responses written outside of the designated areas (we will refer to this as marginalia). In essence, the procedure was changed to guide and direct the edits analysts to establish a "valid" response, if available, applying their judgment when deciding whether or not to key marginalia, instead of rigidly following a prescribed set of rules which in most cases collected unwanted data. Further training was given to the edits analysts in order to hone their discretionary ability in this area and align with Census Bureau policy. After implementing the changes, the team reprocessed every form that had a marginalia anomaly up to the implementation date (where the implementation

[6] The D-1PR(S) was the standard form used in Puerto Rico. It was a Spanish language form distributed using Update/Leave to most households in Puerto Rico and was also sent to those calling to request a questionnaire from Telephone Questionnaire Assistance.

date was defined as when edit analysts were trained and had time to get used to the new rules) (Lockheed Martin, 2010d).

Due to the "subjective" nature of marginalia, accuracy data will not provide much insight into whether or not the objective of marginalia reprocessing was met. This was the first time that the Census Bureau attempted to capture and process this type of information, so it should come as no surprise that we did not know what to expect. One way to examine the effectiveness of this approach is to sample the data and ensure the expected marginalia was being captured. The Data Quality Team performed this analysis and the data met the expected results. Another method was to investigate whether the consistent capturing of the marginalia had an effect on the number of person panels captured. The Data Quality Integrated Product Team investigation and operational observations resulted in the reprocessing of questionnaires because of marginalia. A total of 1,298,638 forms were reprocessed for marginalia (Lockheed Martin, 2010d). The reprocessing changed the person panel count for these forms by 16.88 percent (Lockheed Martin, 2010d). The total population for these forms was reduced by 11.23 percent as a result of marginalia reprocessing (Lockheed Martin, 2010d). This was mainly the result of deceased or duplicate persons.

- Multi Mark Reprocessing: The Data Quality Team noted that there was an anomaly for race and Hispanic origin fields. Responses were causing DRIS to slightly overstate the actual frequency of multi-race or multi-Hispanic origin for conditions where OMR software detected multiple marks. After implementing a complete solution, DRIS and the Census Bureau stakeholders made a decision to reprocess every form identified with multiple checks in a single question, regardless of condition. This was conducted manually with Subject Matter Experts primarily from the Census Bureau's Population Division overseeing the process onsite where all the selected forms were manually keyed under "newly" created procedures (Lockheed Martin, 2010d). The Subject Matter Experts sometimes had difficulty giving consistent direction.

  The Multi Mark Reprocessing yielded a significant change to the Race and Hispanic Origin fields. About 16.14 percent of race responses changed from a multi mark to a single mark. For the Hispanic Origin field, 33.88 percent of the reprocessed responses were changed from a multi mark to a single mark (Lockheed Martin, 2010d).

- A total of 1,464,175 forms were reprocessed either for multiple marks or containing both marginalia and multiple marks (Lockheed Martin, 2010d).

DRIS system engineers were given criteria for a query to be run to pick certain forms for Reprocessing by the Census Bureau stakeholders in the Population Division. For example, system engineers had to write a query that said "Give me all D-1 forms with multiple marks, all

D-1PR(S) forms with a middle initial present, etc." (U.S. Census Bureau, 2011c). Once the query identified the correct forms, there was a workflow management tool that provided the IDs from the query (ibid). Then system engineers created electronic batches of images, with a unique batch ID, including a special character to flag the batch as a reprocessing batch (U.S. Census Bureau, 2011c). System engineers had to then place these forms into the workflow for reprocessing (U.S. Census Bureau, 2011c).

The impact of reprocessing on operations was that DRIS had originally planned to scale down staffing in mid-July (U.S. Census Bureau, 2011c). Instead, DRIS kept analysts and keyers on staff longer (through late August) to accomplish the reprocessing of certain fields on the questionnaire. DRIS expended resources by creating additional training materials specifically related to the reprocessing tasks and then conducted re-training (U.S. Census Bureau, 2011c). During the initial keying for questionnaires, the keyers worked under a performance based pay system. For the reprocessing task, however, DRIS paid keyers at the 100 percent performance level because they were asked to review and key rather than just "key what you see" (U.S. Census Bureau, 2011c). This change was needed because the keyers could not be held to quotas of key depressions per hour when they were being asked to look carefully at each form and examine the context of every response entry presented to them (U.S. Census Bureau, 2011c) similar to the edits clerks.

DRIS used the best overall performing keyers for the reprocessing effort (U.S. Census Bureau, 2011c). These special reprocessing keyers (called edits analysts) could refer the more difficult cases to Subject Matter Experts (SMEs) from the Census Bureau (U.S. Census Bureau, 2011c). Population Division could not supply enough Subject Matter Experts, so Population Division trained and provided oversight to Census Bureau DRIS PMO staff to help perform the SME function (U.S. Census Bureau, 2011c). DRIS carefully monitored schedules because work was scheduled to take place at each Data Capture Center in subsequent time periods, so that the limited team of Headquarters-based SMEs could go to each site in turn to complete the SME portion of the reprocessing work (U.S. Census Bureau, 2011c).

The multi-mark reprocessing effort resulted in three change requests, which when combined together, cost more than $1,150,000 in additional development and labor costs (Lockheed Martin, 2011b). The additional cost of having Subject Matter Experts from the Population Division and Decennial Systems and Contracts Management Office on site at each of the Paper Data Capture Centers to interpret the inconclusive responses during the reprocessing effort was estimated at approximately $88,000 (Latham, 2010). In addition, the reprocessing effort required other added costs in various areas of the DRIS program which cannot be easily quantified, such as the cost of the lost keyer efficiency resulting from applying new processing rules and business procedures during a mature production environment with a very large work force. Many of these costs were absorbed by the DRIS program operating budget utilizing the DRIS budget contingency. A key concept motivating these efforts was the non-monetary costs of error. These efforts were deemed cost effective because of the resulting quality improvement.

If automated data capture was given more consideration earlier in the 2010 Census process and adequate testing was conducted related to understanding how the proposed forms design would affect respondent behavior, the multiple mark issue could have been detected early enough requiring minor adjustments to DRIS and eliminating the need to re-process any of these forms.

One solution which would have required only minor adjustments to DRIS could have been to adjust the check box design to fit the overall capture method, automated and manual. Many of the check box fields were too tightly spaced for automated capture such that the tail of a checkmark could trail into another box.

Had automated capture been considered during tests earlier in the decade, form spacing could have been better optimized for automated capture. If known early, one option could have been to have the Government identified inconsistencies be accommodated by the capture design, but the Government chose to direct the contractor to provide the solution without fully understanding the complexities involved with capturing ambiguous responses.

In the future, the form design process and the chosen capture (automated and manual) method needs to be a fully integrated process involving all stakeholders to ensure the response input and processing requirements can work within the scope of a decennial program. While respondent friendly form design should be the primary consideration, data capture considerations should not be ignored in form design.

In the future, a trade off discussion is needed about the cost and quality benefits of whether to do analysis or key what you see.

The cost of marginalia reprocessing was also covered in these change requests (Muenzer, 2011b). During production we noticed that the edits analysts at all three data capture centers were experiencing a high level of difficulty determining what the intended, detected marginalia responses were and if they were relevant to the question being asked using the existing procedures (Business Rules).

It was not until DRIS was in full production that this discrepancy was understood at a higher level. The integrated DRIS team quickly established a Census Bureau approved updated procedure (Business Rule) directing all three data capture centers on what to capture and what not to capture when presented with these ambiguous responses. Another difficulty was that this change was implemented at all three data capture centers to a full, mature production environment. This evolution occurred as a result of the process of developing the edits subsystem, particularly Consistency Review, which provided a useful platform for drawing attention to these problems. In addition, sizing, staffing, training, and operating the edits subsystem prepared the reprocessing team for addressing the ambiguities encountered.

5.8. Question 8: How did field procedures and materials impact data capture?

- Field Division (FLD) procedures for handling erasures negatively impacted data capture (U.S. Census Bureau, 2011c). As a result of findings from the Fall 2009 Operational Test and Dry Runs, DRIS made a recommendation in December 2009 for FLD to change procedures regarding the handling of erasures, from ensuring that fields were erased completely to instead having FLD draw a line through erased fields. Since FLD procedures had already been printed, kitted, and distributed by this point in time, the recommendation was too late to include in the training materials. However, the instruction was given to FLD to instruct enumerators to draw a line through incomplete erasures, yet this was not done (Gunnison Consulting Group, and ADI, LLC, 2011). The overall OCR accuracy error in DRIS was very low, but a significant portion of the overall OCR accuracy error was due to erasures on the enumerator forms (U.S. Census Bureau, 2011c).

- For GQV, DRIS was tasked with preparing Address Updates (ADDUPs) by applying a set of very complex rules to the data captured from GQV questionnaires. When FLD procedures were not correctly followed for GQV, this sometimes created scenarios that were not envisioned by the ADDUP rule set, resulting in data that DRIS could not deliver as address updates. Data capture operations were unaware there was an ADDUP rule in place that once an ADDUP record was sent to GEO, it could not be resent (U.S. Census Bureau, 2011c).

- The Lessons Learned on Erasures produced by the PDQ contractor indicated that the write-in keying and check-box keying error rates decreased significantly after the business rules change was applied to the handling of erasures in DRIS procedures (Gunnison Consulting Group, and ADI, LLC, 2011a). Table 5.7 shows the magnitude of the change in error rates resulting from the business rules change for erasures on forms.

Table 5.7 DRIS Error Rates With and Without Erasures

Erasure Analysis, D-1(E) Results	Error Rate Excluding Erasures	Error Rate Including Erasures
Write-in Fields	0.48%	0.61%
OCR	0.40%	0.47%
Write-In Keying	0.91%	1.36%
Check-Box Fields	0.01%	0.03%
OMR	0.00%	0.01%
Check-box Keying	0.18%	0.75%

Source: Gunnison Group and ADI, LLC, 2011a

The business rules change was essentially that keyers were instructed to use their judgment to key "intent" under a very clear set of rules, rather than "key what you see." For example, keyers were instructed to pan around the image in an attempt to ascertain if the entire form was lightly written or if the particular response in question is lighter than others on the form, to surmise if they felt it was an erasure. If they felt this was an erasure, they were instructed to press a function key which set an erasure flag in the metadata. This business rule encouraged keyers to hit a function key for suspected erasures. The keyers were instructed to pan around the image and make a determination based on the darkness of the writing across the entire form, using their best judgment, as to whether the form was erased. Keyers were never hired or intended to serve the role of "analyst." That was a function for the edits operators. Trying to change the direction of a large operation after months of consistent direction was very difficult.

The use of color imaging allowed for this level of detailed keyer inspection similar to actually having the paper form at hand for direct examination. The Census Bureau specifically required the use of color imaging in the contract for just this reason. DCS2000 used bi-tonal (black and white) images which did not allow for this level of differentiation. Keyers were instructed not to "Key what you see," but to instead attempt to ascertain the intent of the data on the form and blank out erased text, while keying text they suspected of being lightly written. The use of color imaging also allowed for automation to process forms completed with a wider variety of writing instruments.

- In some instances, FLD procedures were in conflict with data capture rules. For example, the FLD procedure for Targeted Non-sheltered Outdoor Locations during Service Based Enumeration calls for writing names such as "PERSON 1" (U.S. Census Bureau, 2009e), yet Population Division specified that the name field cannot contain numeric values (U.S. Census Bureau, 2011c). The Decennial Management Division did not have a process in place to check these rules against the myriad of procedures. This inconsistency frustrated keyers who were unable to key what was written on the forms even though in this case they were told they had to "key what you see" or risk being charged with a keying error (U.S. Census Bureau, 2011c). Although this was the case, keyers were given a clear set of business rules that instructed them to use Roman numerals in these cases. Some of the problem was they did not clearly understand those rules. The Roman numeral rule was intended for Person Name fields on household based forms, but technically could also be applied to GQE forms. Both the contractor and the Census Bureau should have realized this issue before production.

- The communication of rules for keying was not well coordinated between Field Division (FLD) and Population Division (POP) prior to the release of training materials, resulting in FLD procedures that were in conflict with keying rules. There needs to be more integration between the different areas within FLD and POP, HQP, and the data capture provider. Each response field on each specific questionnaire must have an owner assigned to that field who can speak to the entire lifecycle usage and processing of that field across the entire census enterprise (U.S. Census Bureau, 2011c).

  DMD was the coordinator for keying rules across the different areas including Geography Division, Housing and Household Economic Statistics Division, Field Division and Population Division. A stronger coordination effort could have prevented most of the issues relating to data capture specifications detailed in this assessment.

- The unfolding of forms by the LCO into boxes was a success. This minimized the effort at the Document Preparation stage and allowed forms to flatten out in preparation for scanning which expedited processing of the enumerator forms (U.S. Census Bureau, 2011c).

5.9. Question 9: Were there any requirements which were not clearly defined so that the contractor could provide the intended outcome?

Generally, all requirements were defined such that the contractor provided the intended outcome. However, there were a few instances with high level requirements that were ambiguous or vague by nature and did not give the Census Bureau the intended outcome due to a process that did not include all the appropriate stakeholders. Some of these requirements are described below.

- By agreement with the DRIS PMO, once a Problem Trouble Report (PTR) was opened for Cost and Progress reporting, the Census Bureau would continue to positively acknowledge Cost and Progress files even with defective data (Lockheed Martin, 2011c). In a few instances (such as the errant counts of forms that headquarters processing successfully acknowledged) PTRs were opened. This approach did not give us the intended outcome of having accurate data reported to the Cost and Progress system on a daily basis during most of the production period. However, once a fix to the system was implemented we did receive accurate data near the end of operations and final data was accurate.

- To facilitate processing, DRIS used an architectural design feature called fail queues (Lockheed Martin, 2007). This allowed problem cases to be placed aside (in a fail queue) so that production work was not delayed. As a result of this design, records rejected by Headquarters Processing (HQP) were placed into such a fail queue and not dealt with immediately. This resulted in all of the problem cases previously rejected by HQP to be transmitted to HQP one time near the end of production once a fix was applied (U.S. Census Bureau, 2011c).

  The original requirements for priority processing were not clearly defined by the Census Bureau to include the prioritization of rejected records. Over 10,000 such records were transmitted to headquarters at the end of operations (Lockheed Martin, 2010d). The result is that the intended outcome, of having rejected records processed with their original priority, was not achieved for this small number of records (Lockheed Martin, 2010d).

There was a very complicated priority matrix placed in the contract that created a requirement that was difficult to interpret (U.S. Census Bureau, 2009a). The matrix outlined priorities of forms derived by form types, dates of processing, dates of field operations, and a number of other parameters. The contractor continually sought clarification of the meaning of this priority matrix (U.S. Census Bureau, 2009a) and its place within the contract with other competing similar requirements. Essentially all requirements for priority processing were met.

- Handling and processing of GQV proved to be fundamentally different than processing of household-based questionnaires (U.S. Census Bureau, 2011c). GQV is a validation instrument, not a household based questionnaire. The initial DRIS Request for Proposals did not call out GQV for special processing requirements so that GQV was treated just like any other questionnaire. However, the basic requirements were different. For example, DRIS was required to produce Address Updates by applying a complex set of

rules to captured GQV data, and then transmit these ADDUPs to Geography Division (GEO). Normally, GEO is responsible for producing ADDUPs. For household questionnaires, DRIS simply normalized and integrated these data into a consistent format and transmitted the data to HQP, who in turn, processed the data.

- DRIS used the same reporting system for GQV that they intended to roll out for 2010 household questionnaire mail returns and enumerator questionnaires. This resulted in no progress information sent to Cost & Progress (C&P) on Address Update (ADDUP) creation and all progress reports were on paper processing only and did not show progress or status of ADDUP creation or delivery. The Census Bureau did not request separate reports for the progress of ADDUP creation.

- The specification given to DRIS for preparing the GQV ADDUP was an incomplete rule set that did not adequately cover every eventuality encountered in production. The requirements for handling of these off path scenarios were not clearly defined up front. When the rules set for programming were being developed by DRIS, the Geography Division (GEO) resources were not available because of the priority to complete production work on Address Canvassing. DRIS had already done significant development work based on the preliminary rule set before GEO experts could review the solution. In addition, GEO did not anticipate and plan for the number of exceptions or off path data received on the completed GQV questionnaires which meant that DRIS could not program for those exceptions before production began. A daily (Monday-Friday) Technical Review Board discussion dealt with exceptions on a case-by-case basis during production.

- The delay of the Address Canvassing operation subsequently compressed the GQV operational and data capture timeframe. The timeframe was also compressed because all of the OLQ data collection address update information had to be delivered in order for Headquarters Processing to meet the enumeration universe delivery schedule (Williams, Barrett, 2011). Ultimately, the timeframe was compressed from six weeks to four weeks. The impact to DRIS was there were changes to the processing flow for ADDUP deliveries. Instead of holding everything until the end, DRIS was required to process ADDUP deliveries on a flow basis to allow GEO additional time to iteratively process the ADDUPs. This approach sometimes resulted in incomplete ADDUPs or data that never got transmitted to GEO because of items from the same Assignment Area (AA) that did not arrive together or get processed together.

- GEO could only process each AA once, so if an incomplete AA was sent and materials later showed up from that AA, the subsequent materials may have been lost. For more information on the cost impact of these changes, see the description of DRIS Change

Requests 270 and 271 described in the response to Question 10. This may result in other potential downstream program impacts, including potential increase to the Count Review or Count Question Resolution workloads. Before a contractor is allowed to unilaterally request a schedule delay (such as the Address Canvassing delay), all downstream impacts must first be considered (U.S. Census Bureau, 2011c).

- Population Division (POP) specified global keying rules. These were intended primarily for housing unit based enumeration questionnaires, but some of these rules created issues for GQV (U.S. Census Bureau, 2011c). For example, POP had a global keying rule that said "Do not key refusal explanations such as 'N/A'" (U.S. Census Bureau, 2009a). The presence of information such as "N/A" in the unit designator field is critical to GEO because it provides additional meaningful information that can be used in the ADDUP creation and further downstream processing (U.S. Census Bureau, 2011c).

- There was a rule not to key fractions (U.S. Census Bureau, 2009a). Fractions are a perfectly legitimate entry in an address field. The Paper Integrated Product Team found as a key lessons learned that every data field/item on every questionnaire should have a division owner identified for that field who can develop keying rules and data requirements specific to that item in conjunction with the capture method (U.S. Census Bureau, 2011c). Regular meetings between specific questionnaire data field stakeholders were not held which resulted in inconsistencies between capture and keying of specific data fields.

- In terms of the paper destruction, the Census Integration Group (CIG) developed guidance that paper materials could be destroyed once acknowledged by HQ Processing, but did not initially give specific guidance for GQV forms whose data do not go to HQ Processing, but instead go to GEO. As a result, NPC had to create additional warehousing space for materials that had not yet been authorized for destruction. NPC also ran short on processing carts. Once forms were destroyed, this freed up additional processing carts. NPC had to establish procedures for staging materials for destruction. DRIS had to add steps to the workflow and modify the warehousing application to include flags about which materials were authorized for destruction.

- DRIS keyed fields on enumerator forms in 2010 that were not keyed in Census 2000, and modeled form throughput rates and did not account for keying of auxiliary information that was not keyed in DCS2000. In the future, the Census Bureau should apply an estimated percent filled in metric to all fields indicating what the expectations are for each field to be filled in to allow for better throughput modeling. Enumerator forms initially had lower keying throughput than DRIS modeled for 2010 because of the capture of the auxiliary information. The fact that there was more data than originally modeled,

coupled with PBOCS problems, resulted in a growing concern that the data from enumerator forms would not be captured in a timely enough manner to allow Headquarters Processing sufficient time for end of NRFU reconciliation.

The rate at which the write-in fields had to be sent to a keyer in 2010 was significantly lower than during Census 2000 data capture. The improvement in automated processing in the 2010 Census resulted in 11.76 percent of the write-in fields going to keyers, well under the 16.13 percent that went to keyers in Census 2000. By reducing the keying workload by 4.37 percent from the 2000 workload, DRIS data capture estimated a savings of approximately thirty six million dollars (Coon, 2010).

Once the auxiliary information was recognized as the cause of the slower than modeled form throughput, Census Bureau managers attempted to identify downstream users of this information. At the time, no downstream users of the proxy information were identified. It was later discovered that due to a proxy interview rate much higher than in Census 2000, the Census Coverage Measurement operation could have used this information (U.S. Census Bureau, 2011c).

DRIS Change Request #916 was issued at a cost of $1,800 to address this concern. The CR implementation cost a whole day of down time at each of the data capture sites. The change also compromised the status reporting both internal to DRIS and to the Cost and Progress Reporting (for keying metrics reports not used in the creation of this assessment) because the change resulted in reports no longer accurately reflecting the keying work that had been performed.

The DRIS system is very complex. The system essentially distributes the data capture fields on each form to different subsystems sending some fields to numeric keyers, others to alpha keyers, and still others to alphanumeric keyers or checkbox keyers as part of a complex workflow process. The system must then integrate all these components together once the fields are keyed. Implementing a new process with forms in various stages of the workflow and then trying to put forms back together that used different configurations of the same software, created a high risk, complex set of modification and coordination issues for DRIS engineers.

The lessons that can be taken away from this are that all downstream users must be clearly identified up-front, and take full ownership of their questionnaire content requirements, and that the Census Bureau should give the contractor adequate operational information and time to impact change requests. The scope of the initial request was to change the requirements on an identified field in the middle of processing. So, there were two problems, the change and the timing. If this requirement was known in the beginning, the Census Bureau could have required that the contractor design all the capture fields to be configurable three ways, either “key only”, OCR only or full capture eliminating the difficulty and risk.

- There were a number of unplanned ad-hoc reporting requirements. To accommodate the Director's late add mailing for housing units that were not on the original address list, there was a request to report weekly the number of these questionnaires checked in. This required that DRIS run a database query to produce a count of these records. There was also a request for on-site Census Bureau representatives or Bureau staff at NPC to provide a weekly manual tally of the number of non-official Census forms containing response data or correspondence. These unplanned requirements took additional effort, but were ultimately accommodated.

- For Group Quarters Enumeration (GQE), several large Group Quarters (GQ) facilities offered to write a program to run Individual Census Reports (ICRs) through a computer printer and imprint them with the information contained on administrative lists (U.S. Census Bureau, 2011c). DRIS had no requirement to capture computer printed responses (U.S. Census Bureau, 2011c). As a result, FLD staff had to transcribe GQE information from administrative lists to manually fill out a separate ICR for each GQ inhabitant (U.S. Census Bureau, 2011c). This introduced additional potential for error into the process (for both transcription error and data capture error) and was tremendously inefficient. In the future, the Census Bureau should consider adding a requirement to accept both computer-generated and manual responses using administrative records (U.S. Census Bureau, 2011c).

5.10. Question 10: What was the impact to paper data capture of late and or ad hoc requirements presented by the Census Bureau to the DRIS contract?

Since all deliverables were provided on time and all service level agreements were met, the only way that the impact of change can be easily quantified is by examining the cost impact of change requests on the DRIS program. Overall, the DRIS projects an Estimate at Completion (EAC) under the total budget allocation. However, change requests added considerable dollars to the total contract costs and contract value.

The original contract award for the DRIS program was approximately $483.2 Million in October 2005 (Lockheed Martin, 2009c). This original contract award only funded the first phase of the contract which covered systems development for Dress Rehearsal (Lockheed Martin, 2009c). There was always a plan to have the contractor separately re-bid subsequent phases of the contract.

In February 2006, an update to Phase I reallocated contract value between Phase I and Phase II of the contract, but kept the overall contract value at approximately $483.2 Million (Lockheed Martin, 2009c). An update to Phase II in February of 2007 increased the total contract value to approximately $562.5 Million (Lockheed Martin, 2009c). This was due to the reallocation of Phase I contract scope and to collaboration between Lockheed Martin and the Census Bureau

regarding the true scope of the contract (Lockheed Martin, 2009c), as well as volume changes due to the paper NRFU and a $26 million increase in the Service Contracting Act Labor rates.

Other Contract change proposals for the Jobs Line, GQV, NRFU Latency, Electronic Fingerprinting, and the American Recovery and Reinvestment Act funding for Coverage Followup resulted in approximately $314 million in additional increases to contract value (Lockheed Martin, 2009c).

By July 2009, the contract value grew through change control to over $867.6 million (Lockheed Martin, 2009c). The final contract value as of February 2011 was $1,019,134,529 (Pentercs, 2011). The largest share of the increase was due to an increase in projected workloads and scope changes on the contract. Some of this cost increase was driven by changes to telephony requirements which are out of the scope of this assessment.

The costs driven by paper related changes over one million dollars each are addressed here.

- The DRIS contract modification pertaining to the change from electronic transmission of data from the handheld devices to paper data capture of Non Response Followup (NRFU) cases added $137,000,000 in contract value to the DRIS contract (Lockheed Martin, 2008b). The funding supported the changes necessary to: extend the full production schedule for two months, purchase increased system capacity, including extending maintenance contracts, site support, and software licenses, and in maintaining resources for the paper data capture center operational support to capture over 63,500,000 additional paper questionnaires (U.S. Census Bureau, 2011b) rather than receiving electronic transfer of data from handheld devices (Lockheed Martin, 2008b).

- Over seven paper and data integration related contract Change Requests (CRs) with a price tag of over one million dollars each were issued during the DRIS program (Lockheed Martin, 2011b)

- The following DRIS change requests that impacted paper data capture had a cost over $1,000,000 each.

- DRIS Change Request #270: This change request compressed the GQV processing timeframe at an added cost of $4,900,000. DRIS Change Request #271 This change request added additional Address Update (ADDUP) preparation tasks to DRIS at an added cost of $1,300,000. Together, these two change requests impacted ADDUP creation and led to some of the issues described in the response to Question 9.

- DRIS Change Request #273. This Change request was for the Fiscal Year 2008 costs associated with preparing the Rough Order of Magnitude estimate for conducting a paper Nonresponse Followup Operation, and added $1,042,901 in costs. This Change Request

ultimately led to the $137,000,000 contract modification described above. As a result of moving from handheld computers to paper based Enumerator operations, over 63.5 million enumerator questionnaires were added to DRIS data capture workloads (U.S. Census Bureau, 2011b).

- DRIS Change Request #295. This change request was for the DRIS contractor to build out and fit-up the National Processing Center to serve as a data capture site, and added a cost of $1,539,302.

- DRIS Change Request #344. This change request was needed to prioritize NRFU scanning to accommodate MaRCS. This CR added a cost of $24,330,650. Some of the impacts resulting from this change request are described further below.

- DRIS Change Request #522. This change request was to provide all DRIS images to the Census Image Retrieval Application, at a cost of $3,138,816.

- The Data Capture requirements for the Evaluation, Experiments, and Assessments program were not known until late in the development cycle for the Paper Data Capture subsystem and the Workflow Control and Management subsystem. This resulted in a contract modification totaling $5,000,000. (Lockheed Martin, 2011b). This is because the requirements for the Evaluations, Experiments, and Assessments program were not known up front.

- The DRIS data capture subsystem and the Workflow Control and Management subsystem were impacted by the move from using handheld computing devices to paper for field operations (Lockheed Martin, 2011b). This change resulted in a CR which (when adjusted for actual cost) cost $17.3 million dollars (DRIS Change Requests #723 and #926) to make program and system changes and to support the extension of data capture operations in the Paper Data Centers (Lockheed Martin, 2011b). Originally, the paper operations were supposed end with the completion of capture of the mail returns and the Workflow and Control subsystem would continue to receive electronic data for the field follow-up operation (Lockheed Martin, 2011b). Before this change request, Vacant Delete Check was only intended to be processed at NPC (Lockheed Martin, 2011b). As a result of this CR, paper data capture operations were extended at Baltimore and Phoenix from May 31st through August (Lockheed Martin, 2011b).

- The change from handheld computing devices also impacted the Matching Review and Coding System (MaRCS) which was developed for automated field data collection reinterview. When the decision was made to go to paper for field data collection DRIS had to purchase more scanning equipment and modify the DRIS paper data capture

architecture and priority workflow to meet the requirement for a 10-day turnaround time from enumerator questionnaire check in and scanning by DRIS to provide data to MaRCS for their Quality Assurance (QA) program. The cost of the change for new equipment and the modification of the DRIS architecture was $24,330,540 (Lockheed Martin, 2011b). The change also meant that DRIS would send raw data capture data to MaRCS rather than the data captured data that was processed through all of the data capture processing subsystems after scanning. Ultimately, both the raw data and the final data with data capture QA steps applied were sent, but because of timing considerations, MaRCS had to get the raw data initially.

5.11. Question 11: How did the addition of late requirements covered within the existing contract scope impact budget and resource scheduling?

The DRIS program was established anticipating a certain level of late requirements. We anticipated more late requirements than actually occurred. In anticipation of this change, money was placed on contract as a management reserve. All of the management reserve was not spent, and these unspent funds were returned to the government.

The DRIS contract was segmented into three phases and there were baselines within each phase. Phase I of the DRIS contract was early planning and included DRIS data capture of the 2008 Dress Rehearsal. Phase II of the contract was essentially 2010 Production Operations. Phase III of the DRIS contract covered archiving and contract closeout activities (U.S. Census Bureau, 2005).

This phased approach allowed requirements to be decomposed with each new phase and base lined, enabling more accurate scheduling, cost, and performance estimates. This phase approached allowed for the latest empirical results to be considered in iterative requirements development (U.S. Census Bureau, 2005). Separating the contract into phases and baselines enabled the government to decompose and refine requirements during the development of DRIS. This resulted in adequate budget and resource scheduling.

5.12. Question 12: Were the Census Bureau deliverables to DRIS timely and complete? If not, what was the effect?

With the exception of printed copies of experimental forms, the Census Bureau deliverables to DRIS were timely and complete, although the contractor's expectations were an earlier delivery date on average. Final printed copies of experimental forms were not delivered to DRIS until after data capture had already begun, resulting in late starts to the forms definition and testing process.

This occurred because the low volume of these forms did not justify the cost of printing earlier, and these forms were printed "just in time" for the mailout. This was also coupled with a very late content finalization, in the fall of 2009, for these forms. These forms could not be designed until after content was finalized in the fall of 2009. This introduced significant risk to processing of these form types and added $287,120 to the cost of DRIS (Lockheed Martin, 2011b).

One effect of this late delivery of experimental forms is that DRIS had to make last minute changes to the sort plan for these forms. As a result of sort plan changes, DRIS reported that over 8,000 of these experimental forms may have missed the 48 hour check-in deadline, placing additional limitations on the Deadline Messaging experiment (MacDonald, 2010). Another difficulty introduced was that operations had to manually set these forms aside after check in for later processing until the scanning templates were patched to the site. If a form went to scanning prior to this patch, they would fail and need to be reprocessed.

While these final printed forms were not delivered to DRIS late because of the deliberate schedule change, DRIS had downstream dependencies on receipt of the printed forms. In the future, schedules need to be electronic and integrated to eliminate discrepancies. DRIS had to manage to these later dates. This change resulted in an additional cost of $301,200 (Lockheed Martin, 2011b).

A special team (Form Design IPT) was set up to coordinate issues with form deliveries and the forms delivery schedule and managed forms design issues affecting data capture. DRIS prepared a color-coded chart that indicated the status of each form weekly. This was an effective management tool to manage form deliveries of draft and final PDFs as well as delivery of first-offs or prior to production print samples of each form. These helped ensure that deliverables were sufficiently complete (U.S. Census Bureau, 2010b).

5.13. Question 13: Were the deliverables from the contractor to the Census Bureau timely and complete?

All paper data capture contract deliverables from DRIS to the Census Bureau were on or ahead of schedule and considered to be sufficiently complete. The contractor implemented a thorough Document Coordination and Approval (DCA) process. As a part of the DCA process, stakeholders were engaged throughout the full life cycle of each deliverable's creation. The DCA process helped ensure that deliverables were complete (U.S. Census Bureau, 2011c).

While not an official contract deliverable, DRIS prepared a set of paper "Forms Design Guidelines," which had spacing guidelines and forms layout guidelines for optimal data capture. These guidelines were first issued by Lockheed Martin in March 2006 and provided to the Census Bureau at that time. However, Lockheed Martin revised the guidelines nine times through October 2010 (Muenzer, 2011c). The changes during the forms design timeframe impacted the ability of the Administrative and Customer Services Division to finalize the design of questionnaires for data capture (Muenzer, 2011c).

5.14. Question 14: What were the impacts of changes on interface management—what worked and what did not?

During GQV, the interface between the Field Data Collection Automation system and DRIS worked well (U.S. Census Bureau, 2011c). The following describes each of the interfaces between DRIS and Paper Based Operations (PBO) and how these interfaces had to be changed during production to account for PBOCS issues (adapted from Lockheed Marin, 2011a):

PBO to DRIS Notification of box shipment:

Box notifications were sent at the beginning of field enumeration production. However, PBO soon encountered capacity issues that limited the number of concurrent workstations that each LCO could support. Because PBOCS was not used for shipments as originally designed, the only inventory control function that was comprehensively performed was the checking out of forms from the LCO to the field. Therefore, DRIS did not reliably receive box shipment notifications. DRIS received shipment notifications for only ten percent to twenty percent of boxes from LCOs. This severely limited the inventory control of boxes between the two systems that DRIS had planned to use to inform their operational staffing. The tracking system of FedEx was used as the de facto inventory control mechanism for DRIS (Lockheed Marin, 2011a).

Linkage of Group Quarters Enumeration (GQE) forms to Group Quarters Identification Number (GQID):

PBO was responsible for creating electronic linkage fields during check-in at LCOs. In addition, the field staff was directed to manually apply the GQID label on each GQ form. Since the individual GQE questionnaires have no geographic information about the GQ location, it is this linkage that electronically associates (or links) each GQE questionnaire to the specific GQ location. However, neither of these linking actions were consistently performed (Lockheed Marin, 2011a). DRIS estimated that approximately seventy-five percent of GQE forms were linked to GQIDs based on linkage files from PBO. When this interface became unreliable, NPC's Automated Tracking and Control system (ATAC) was used to create linkage files for the remaining twenty-five percent (Lockheed Marin, 2011a). However, approximately 88,000 forms required even more special handling (Lockheed Marin, 2011a). These forms were unlinked electronically because they were missing hand-written GQIDs. To link these forms, a quality inspection step was added within DRIS operations to inspect and manually transcribe, if necessary, the GQID from the D-352 Group Quarters Enumeration Record to the back of each GQE form. Following this action, a GQ linkage tool created by DRIS was used to establish the necessary linkage. Of the 88,000 unlinked GQE forms, approximately 70,000 were linked using this tool. The remaining 18,000 were linked by a manual analysis (Lockheed Marin, 2011a).

Linkage of enumerator continuation forms to parent forms:

The enumerator questionnaire only has space for five people. For large households containing more than five people, continuation forms must be completed. The continuation forms contain no geographic information about the household and they are electronically associated (linked) to the main (parent) enumerator form for that household. The same capacity issues of PBO rendered this linkage information unreliable. Many continuation forms were received from the LCOs before the linkage information was received electronically. A special linking application was created by DRIS, and used to periodically search for the linking data necessary to associate continuation forms that were previously unlinked.

DRIS to PBO Notification of box receipt:

DRIS transmitted receipt notifications to PBO for all boxes received. However, due to the limitations of PBO, these data were not used. Very few acknowledgments were received from PBO in response to these notifications.

Notification of form receipt:

DRIS transmitted receipt notifications to PBO of all forms received from the field. During most of the production period, no acknowledgements were received in response, indicating that they were not being processed.

As a result of PBOCS' development issues, DRIS was directed by a change request to develop a contingency to receive Enumerator Continuation Form linkage data for NRFU Vacant Delete Check in the event that PBOCS would be unable to provide these data (Lockheed Martin, 2011c). This contingency was not utilized (Lockheed Marin, 2011a). Had we needed to use it, it would have leveraged the existing Cost and Progress interface with DRIS to avoid creating a new System Data Exchange (Lockheed Marin, 2011a). Cost and Progress was the only DRIS interface partner that also had connections to AMSD, who developed the contingency system for VDC Linkage information (Lockheed Marin, 2011a).

The original baseline for DRIS contained funds for some end of program reconciliation tasks. However, change requests totaling over $577,456 were issued to account for additional end of program reconciliation costs resulting from PBOCS issues (Lockheed Martin, 2011b). Much of this was due to special queries that were required to accomplish linking Enumerator continuation forms to parent forms for large households.

DRIS had over nine interface partners with over seventy system data exchanges with other systems from the start of GQV processing in 2009 through the post-production delivery of images to the Census Image Retrieval Application (CIRA) in 2011 (see the table in Appendix E for a listing). This was seen as too many interfaces (U.S. Census Bureau, 2011e). Some of the interfaces received very similar products/files to those going to other interfaces (U.S. Census

Bureau, 2011c). It would have been better to streamline the data exchanges into fewer deliveries and let end users pick and choose what data they needed rather than having DRIS prepare custom deliverables for each user that were only slightly different from one another reducing the DRIS overhead (U.S. Census Bureau, 2011c).

The defined process for Interconnection Security agreements was seen as a success. The management of External Interface Control Documents (EICDs) overall was successful, but was a tremendous effort (U.S. Census Bureau, 2011c). The DSSD EICD contained over 21 system data exchanges alone (U.S. Census Bureau, 2011c). Many of the interface partners in DSSD did not even know what their system was going to look like or what they even needed when the EICD was established, because DRIS's development schedule was ahead of that for DSSD (U.S. Census Bureau, 2011c).

The EICD process was not well defined (U.S. Census Bureau, 2011c). Some stakeholders required a signature on the EICD, while others required a Memorandum of Understanding between interface partners before the EICD could be signed (U.S. Census Bureau, 2011c). In some cases, the EICD became the vehicle to define and collect requirements, rather than reflect the existing requirements which was beyond the scope originally intended for the EICD (U.S. Census Bureau, 2011c).

The Product Services Message Queue (PSMQ) was the mechanism for handling all interfaces between DRIS and Census Bureau Headquarters (U.S. Census Bureau, 2011c). The centralized approach of interfaces managed through one place (PSMQ) worked well (U.S. Census Bureau, 2011c). PMSQ was developed in house which required DRIS to participate in much more testing than planned and more than would have been required if Headquarters had used a commercial off the shelf product such as Oracle Message Queue (U.S. Census Bureau, 2011c).

5.15. Question 15: Were the testing methodologies used to test systems and operations sufficient to meet requirements?

The testing methodologies were both adequate and sufficient and with all of the appropriate stakeholder involvement for Phase I, Phase II, and Phase III. DRIS had a working methodology called the Test Evaluation Master Plan (TEMP), as well as, all lower level testing plans in place for adequate testing within each of the Segment Channels (U.S. Census Bureau, 2011c). The original draft of the TEMP for the 2008 Census Dress Rehearsal, provided by the contractor, had to be reworked, but this was done prior to any testing (U.S. Census Bureau, 2011c). The TEMP was adequate (U.S. Census Bureau, 2011c). The DRIS contractor modified the TEMP for the 2010 Census. It was reviewed several times, vetted through the Document Coordination and Approval process, and with stakeholder involvement (U.S. Census Bureau, 2011c). The PMO felt all testing documentation was adequate for the Phase II of the DRIS contract for the 2010 Decennial Census (U.S. Census Bureau, 2011c).

Lessons Learned participants determined that a tremendous opportunity was missed by not having the DRIS system certified to accept Title 13 data earlier in the 2010 Census cycle (U.S. Census Bureau, 2011c). There were many rich data sets from mid-decade tests and from the 2008 Census Dress Rehearsal that would have allowed early 2010 Census tests to use real-world data with forms filled out uniquely by many different households and a variety of handwriting styles or writing instruments U.S. Census Bureau, 2011c).

Had DRIS been able to use 2008 Census Dress Rehearsal materials (or other materials from mid-decade tests) for testing in the 2010 Census cycle, including envelopes, many of the issues ultimately experienced when 2010 Census data were encountered could possibly have already been addressed through testing. The Automated Marginalia detection feature only had very limited testing using actual data. There was only limited contextual testing with POP and other subject matter analysts. This was a key functionality of the DRIS system. This is one feature that allowed reprocessing to occur. Marginalia appeared on a number of 2008 Census Dress Rehearsal forms, and yet, went untested prior to going into production for the 2010 Census (U.S. Census Bureau, 2011c).

6. Related Evaluations, Experiments, and/or Assessments

The 2010 Group Quarters Validation (GQV) Assessment
The 2010 Telephone Questionnaire Assistance (TQA) Assessment
The 2010 Coverage Followup (CFU) Assessment
The 2010 Universe Control and Management/Response Processing Systems Assessment
The 2010 Content and Forms Design Assessment
The 2010 Forms Printing and Distribution Assessment
The 2010 Mail Response/Return Rates Assessment
The 2010 Census Nonresponse Followup Reinterview Quality Profile

7. Lessons Learned, Recommendations, and Conclusions

Lessons Learned Summary

Thousands of individual detailed Lessons Learned were documented, by the contractor, the DRIS Program Management Office, and DMD. Below are many of the key lessons learned:

- Any form design process or evaluation, which utilizes a data capture process, needs to be fully integrated with the capture process requirements very early.
- Stakeholders must own all response fields throughout the entire capture process.

- We must establish mature, realistic requirements early in the process. Depending on the change control process and overhead could equal program failure and major cost overruns.

- "Adequate" testing must be completed frequently and early with all stakeholders involved.

- For all 2020 design and processing areas, Census should first establish what "best practices" were used in the past and start from there. This process can avoid decisions that would just lead us back to the "current" best practice due to a lack of communication between programs.

- Make sure that if the contract workload estimates include a contingency that the contractor does not apply their own contingency on top of the Census Bureau's contingency already built into workload estimates.

- Every data capture field on every form needs to be assigned to a specific subject matter expert (SME) who owns the entire life cycle usage and processing of that field. That SME owner must be empowered to resolve disputes between different subject areas over the data capture rules for that field. The SME's should fully understand the "selected" data capture process in the future.

- Reporting requirements should be defined earlier in the data capture system and workflow and control design phase. For the 2010 Census, it would have been better to develop the reports based on a full dress rehearsal and lock the reports for the duration of the decennial census.

- During the creation and updating of user guides, procedures, and training addendums, procedure walkthroughs should be conducted and procedures should be clearly defined, leaving little leeway for misinterpretation by each data capture site. More clarity is needed in procedures to avoid having each site making different decisions and "cheat sheets" leading to inconsistency in operations and the data output. Procedures should be nimble enough to adapt to situations or issues occurring regularly that procedures do not provide guidance for.

- Greater coordination is needed with the USPS to avoid intermixing of different materials that should have been pre-sorted at the post office based on the different Zip + 4 information. NPC did not have this issue due to past experience working with USPS properly configuring "Zip+4" presorting from the USPS.

- Production operations such as GQV should be better coordinated with the timing of Operational Test and Dry Runs (OTDRs) for census production to avoid having the same testing resources spread too thin.

- The Census Bureau must give the contractor more consistent direction on how to handle Title 13 data such as addresses. For example, should paper materials containing address information only be shredded or recycled? For the 2010 Census, the Census Bureau opted for the most conservative approach and shredded everything.

- As with 2010, there needs to be a solid agreement between the data capture Contractor/Service Provider and the Census Bureau managing form design requirements very early. Form design guidelines for automated data capture must be fully agreed upon early in the questionnaire design process. As for DRIS, minimum spacing tolerances along with other capture specific requirements for data capture communicated in the forms design guidelines should be carefully evaluated during design and development and then rigorously tested against a test deck with corresponding truth data available. Consideration should also be given to reducing the amount of unique paper form types processed.

  The paper system must consider:

  - Examining the amount of white space around the check boxes including the box placement and determining what might be captured with various styles of responses.

  - Having a field-by-field capture/design requirement, instead of an overall form requirement.

  - Recognizing the varying importance and response rates of different fields to properly classify them during the design process. This would have allowed low use fields to be closer together than high use fields.

Recommendations:

- The best and most realistic solution for 2020 may be a hybrid solution. The goal should be to maintain ground gained in technology advances and reducing costs. Utilizing the experience and control of an "In House" solution then taking advantage of strong technical advantages available only through private industry. The in-house effort should be to become integrators as well as developers. To accomplish this, the current internal data capture organization will need to establish a new program office with aggressive goals to lead the industry in this area.

- For paper processing, the "cluster" type module architecture was utilized for the 2000 and 2010 Census paper processing capabilities. This type of design has been proven

successful for two decennial size paper capture processing programs known to be among the largest in the world.

- A research effort should be started investigating ways and methods to understand what portions of the DRIS program could have been cut or reduced while still providing the basic functionalities required by Census. The DRIS program should be used as a best practice model for data capture.

- The Bureau needs to establish clear and appropriate data capture requirements as we move forward to 2020 and beyond. Requirements need to meet all the Bureau standards, needs now and in the future. Requirements should not be created to fit the current "in-house" solution or any other limitation but require the selected provider to meet current and future data capture requirements. Solutions need to be "process dependent" demonstrating an open source type of supportability structure, not depending on any one single point of failure reducing risk to the Census.

- The proposed corporate solution for data capture- The newly created corporate solution for all Census data capture programs should follow the same evolution process that Census Printing Services, Network Services, Desktop support, Server support and SAN Storage was "required" to follow. All noted services were required to meet industry "best practices" which Census was required to maintain.

- Form content cannot be developed in a vacuum. When form content is determined, data capture procedures and methods, FLD operational procedures, and the entire lifecycle usage and processing of every content item must be considered before the form is designed. Staff for every step of the process of data collection, from form content determination to form design, to fielding the instrument, to data capture, to processing the data, should be engaged from the beginning of the development process. If data capture needs are not considered during design and content determination, stakeholders must be made aware of the impact to final data capture results.

- Every item on every form should have a clear stakeholder identified who owns the entire lifecycle usage and processing from collection in the field, to data capture, to post processing. This should be done to avoid conflicts between FLD procedures and data capture procedures or processing procedures. Doing so for the 2010 Census could have helped to avoid problems such as those seen with the inability to key numeric values in the name field or to distinguish ambiguous erasures.

- While user friendliness should be the primary consideration in form design, the method of data capture and data capture considerations and guidelines should be taken into account from the beginning. Because data capture was not a test objective in the decade leading up to the 2010 Census (U.S. Census Bureau, 2003b), data capture was not

sufficiently considered early enough in the form design, layout and testing for the 2010 Census, resulting in the multi-mark reprocessing and other rework. A critical element in the process is knowing if the survey needs to conform to a "keyed" process or an "automated" process for data capture. The form design can be impacted greatly by the process required. An example would be the transformation of the GQV booklet from a "key from paper process" to a fully automated process. Census testing leading up to the 2010 Census utilized NPC's legacy "key from image/paper" capture method for most response data which gives more flexibility to the form design process as compared to the DRIS method which was required to use a fully automated capture process to capture response data (OCR/OMR).

- To avoid dramatically over sizing systems, we should have workload estimates determined ahead of time and have documented assumptions indicating how they were derived by the Census Bureau and the contractor. There should be clear information about if these estimates include contingency. If so, how much contingency is included in the estimates, if the design remains the same as 2010? Census 2010 results should be used as the baseline for most predictions.

- More research is needed into marginalia, or other ways that people unexpectedly fill out forms. Capturing the "intent" of a respondent rather than "key what you see" is very difficult, and Population Division should research ways to capture responses from people who do not complete the form in the way that was expected. The telephone number on the questionnaire could be used to follow-up with cases where respondents indicated an ambiguous response with marginalia or multiple marks. Lockheed data indicate that 1,464,175 forms were reprocessed for either marginalia and or multiple marks, and 1,298,638 forms were reprocessed for marginalia only (Lockheed Martin, 2010d), resulting in a total of 2,762,813 forms reprocessed for marginalia and or multiple marks out of a total of 81,209,565 million forms eligible for reprocessing (U.S. Census Bureau, 2011b). In the future we either need to decide to ignore and not capture anything outside of the response fields or give complete guidance to operations on what to capture with a Census SME onsite for direction when inconclusive cases arise.

- Data Quality was a major focus of the DRIS program. An independent group of data quality experts who were contracted separately along with stakeholders from DSSD and POP were essential in identifying the systematic clustering of error in items that eventually were reprocessed in order to repair and improve the data resulting in a more accurate census. An independent data quality system supported by independent experts both outside and inside of the Census Bureau is essential for ensuring data quality and this practice should be continued. In the future, this operation needs to be located at

Census Bureau Headquarters. It is essential that future data capture operations have this type of independent oversight.

- The Census Bureau must establish firm and mature requirements to reduce the amount of change and thereby reduce costs.

- Any future system should have as many "configurable" type requirements as possible. When a change does occur, it will be very easy to change and cost much less than having to re-write software code

- The 2020 Census system should be built as a "future use" system. Not a one-time use.

  The decomposition of requirements must be thoroughly vetted with the appropriate stakeholders.

Conclusion:
Overall, the DRIS program's paper capture and integration components should be considered a success. In spite of challenges from PBOCS and the change from electronic to paper Non-Response Followup, DRIS provided key mitigation strategies that resulted in workarounds so that they completed operations on schedule and under the total budget allocation. The data quality focus of the program both in the development of the DRIS systems and by including independent quality experts ensured that the data captured by DRIS were of high quality. The reprocessing effort was successful in enhancing data capture output by having Subject Matter Experts on site who could attempt to more truly reflect the intent of respondents rather than simply asking keyers to "key what you see." There were issues with discrepancies between data capture and FLD procedures, but DRIS was able to work through these issues.

8. Acknowledgements

The authors of this assessment wish to thank the members of the Paper Integrated Product Team, the Data Quality Integrated Product Team, the Workflow Control and Management team, and all those who participated in the lessons learned sessions and other contract meetings. Special thanks go to DRIS Program Management Office (PMO) staff, which reviewed this assessment for accuracy and completeness and provided comments. DRIS PMO staff also provided much of the data that are contained in this assessment, for which we are eternally grateful. Thanks goes to Janice Pentercs for her assistance providing cost data. Thanks also goes to Ray Muenzer for helping write many of the technical descriptions in this report.

9. References

Cardella, Michael A., Jr, (2010), "Reprocessing Description version 4," Word Document, September 13, 2010.

Coon, David, (2010), “Re: Op Press Briefing - data capture metrics,” electronic mail message, October 29, 2010.

Ciango, Andrew, (2010), “Change Request: Delete the D-21PR Military Census Report (MCR) and D-23PR Shipboard Census Report (SCR) forms from the 2010 Census GQ Program,” Word Document, February 2009.

DeMaio, Theresa, Jennifer Beck, and Laurel Schwede, (2008), Report on Cognitive Pretesting of the Census 2010 Mailing Package Materials, Study Series (Survey Methodology #2008-12), October 20, 2008.

Department of Commerce Office of Inspector General, (2011), “2010 Census: Quarterly Report to Congress,” Final Report No. OIG-11-017-I, January 18, 2011.

Goerman, Patricia et al., (2007), “Census Bilingual Questionnaire Research Final Round 2 Report,” STUDY SERIES *(Survey Methodology #2007-27)*, October 21, 2007.

Gunnison Consulting Group, Inc. and ADI, LLC, (2011), "Production Data Quality Final Report on DRIS 2010 Data Capture Quality," First Draft, Version 1.16, March 31, 2011.

------ (2011a), “Production Data Quality (PDQ) Executive Report on DRIS 2010 Data Capture Quality,” Version 2.0, September 27, 2011.

Largent, Deborah, (2010), “FY 2010 - Outstanding Unfunded Requirements as of April 29, 2010,” PDF File, April 29, 2010.

Lockheed Martin, (2007), "Census Practice System Administration Subsystem Fail Queue GUI Component Detail Design," Version 1.1, April 3, 2007.

------ (2008a), “Decennial Response Integration System (DRIS) Call Center Operations DRAFT,**”** Draft, Version 7.0, October 13, 2008.

------ (2008b), “Contract Change Proposal #17,” November 2008.

------ (2008c), “Paper NRFU ROM (FY 09-10),” April 30, 2008

------ (2008d), "Decennial Response Integration System (DRIS) Operations Plan," First Issue, November 17, 2008.

------ (2009a), “2010 Census Forms list as of 7-17-08a with Site Distribution,” Excel Spreadsheet, April 9, 2009.

------ (2009b), “Decennial Response Integration System (DRIS) Census 2010, Phase II, Baseline 3 Proposal, Volume 5 Management, Solution, and Operational Considerations.” *Procurement Sensitive*. June 17, 2009.

------ (2009c), “DRIS Executive Review,” Power Point Presentation, July 2009.

------ (2010b), “SLAP Summary Operational Contingency Log.” May, 2010.

------ (2010d), “Reprocessing Summary,” PowerPoint presentation. September 21, 2010.

------ (2010), “Census Practice Form Design Guidelines Version 6.0,” September 27, 2010.

------ (2010c), “Final CNP Delivery –Ed.xls, Suppressed C&P Values,” Excel Spreadsheet. October 22, 2010.

------ (2011c), “External Interfaces Control Document (EICD) Decennial Response Integration System (DRIS) To Cost and Progress System (C&P),” January 25, 2011.

------ (2011a), “Decennial Response Integration System (DRIS)
Data Quality Management Plan and Production Results,” March 30, 2011.

------ (2011b), “Decennial Response Integration System (DRIS) RAT PACK SLAP CR TRACKER,” Excel Spreadsheet. April 4, 2011.

MacDonald, William H. (2010), “Re” cp_data_20100318_1.dmp - successful Delivery,” electronic mail message, March 19, 2010.

Muenzer, Raymond (2011a), “# of Clusters,” electronic mail message, January 13, 2011.

------ (2011b), “Writeup of reprocessing cost” electronic mail message, August 5, 2011.

------ (2011c), “Action Items” electronic mail message, September 13, 2011.

Pentercs, Janice L. (2011), “Re: Fw: DRIS Costs,” electronic mail message, August 5, 2011.

U.S. Census Bureau, (2001a), ”Program Master Plan: Census 2000 Data Capture Systems and Operations,” Census 2000 Informational Memorandum, No.107, June 21, 2001.

------ (2001b), “Census 2000 Telephone Questionnaire Assistance Program Master Plan (PMP),” Census 2000 Informational Memorandum, No. 111, August 14, 2001.

------ (2003a), “Census 2000 Informational Memorandum No. 135,” February 19, 2003.

------ (2003b), “2004 Census Test Operational Plan,” September 29, 2003.

------ (2004), “Cost and Progress Report: 2004 Census Test Data Capture,” September 2004.

------ (2005), “Decennial Response Integration System Contract, Section C,” December 2005.

------ (2006a), “2006 Census Test Objectives and Research Questions, Ranked by 2010 Research and Development Planning Groups,” Draft, February 16, 2005.

------ (2006b), “2010 CENSUS TEST MEMORANDA SERIES Chapter: 2005 National Census Test No. 19,” March 2, 2006.

------ (2006c),“2010 CENSUS PLANNING MEMORANDA SERIES No. 47,” March 3, 2006.

------ (2006d),“2005 National Census Test Self Response Options Analysis,” November 20, 2006.

------ (2008a), “2010 CENSUS TEST MEMORANDA SERIES Chapter: 2005 National Census Test No. 30,” June 26, 2008.

------ (2008b), “2010 DECENNIAL CENSUS PROGRAM DECISION MEMORANDUM SERIES No. 23,” May 8, 2008.

------ (2008c), “NRFU Latency Issue Description and Impact,” October 22, 2008.

------ (2009a), “Decennial Response Integration System Contract,” May 11, 2009.

------ (2009b), “DRIS Paper Questionnaire Priority Processing, by Date, for the DRIS Data Capture Operations,” February 4, 2009.

------ (2009d), “CIG UAA Issue Paper,” July 31, 2009.

------ (2009c), “Keying Instructions At The Questionnaire Level,” September 21, 2009.

------ (2009e), “D-569.12, GQE Enumerator Manual” Chapter 5, p. 5-4.

------ (2010a), “2010 CENSUS INFORMATION MEMORANDA SERIES No. 2 (Reissue 2),” July 7, 2010.

------ (2010b) "Content & Forms Design IPT 2010 Census Content & Form Design Process Lessons Learned, Best Practices, and Recommendations for 2020," Session 1 Meeting Minutes, January 6, 2010.

------ (2011a), "Cost and Progress Report: Data Capture Backlog: GQV." April 12, 2011.

------ (2011b). "Cost and Progress Report: Cumulative Data Capture Progress." April 12, 2011.

------ (2011c). "DRIS Data Capture and Integration Lessons Learned" March 28, 2011.

------ (2011d), "Cost and Progress Report: 2010 UAA Check-In Status by DRIS Data Capture Center," April 12, 2011.

------ (2011e), “PMO WCM Lessons Learned” Excel Spreadsheet, January 21, 2011.

------ (2011f), “USPS/Census Joint Committee Meeting Consolidated Comments Lessons Learned,” April 20, 2011.

Williams, Jean, and Dianne Barrett, (2011), “2010 Group Quarters Validation Operation Assessment Report.” Briefing Draft Report, September 26, 2011.

APPENDIX A: Comparison of Estimated and Actual data capture workloads

Table A1: 2010 Workload Comparison

Form[7]	Estimated Capture Workload	Total Checked in to Data Capture through 9/7/2010	Percent of Expected Workload Checked in	NOTES
D-1*	92,000,000	82,929,578	90.14%	*Estimates were never revised downward to account for change in the Replacement Mail strategy; this initial 92,000,000 estimate included 10,000,000 bilingual forms.
D-1(U/L)	405,675	271,535	66.93%	
D-1(E/S)	10,670,000	7,273,757	68.17%	
D-1(C)	136,000	1,535	1.13%	
D-1(K)	68,000	1,417	2.08%	
D-1(R)	68,000	558	0.82%	
D-1(S)	646,000	55,933	8.66%	
D-1(V)	68,000	1,180	1.74%	
D-10	550,000	681,320	123.88%	
D-10(C)	300,000	10,871	3.62%	
D-10(K)	300,000	4,680	1.56%	
D-10(R)	300,000	2,071	0.69%	
D-10(S)	500,000	71,246	14.25%	
D-10(V)	300,000	3,335	1.11%	

[7] See Appendix B for explanation of Form Types.

Form[7]	Estimated Capture Workload	Total Checked in to Data Capture through 9/7/2010	Percent of Expected Workload Checked in	NOTES
D-1(E)	60,000,000*	57,311,583	97.23%	*Estimate includes D-1(E)X1 and D-1(E)X2
D-1(E)SUPP	3,039,777	1,858,977	61.16%	
D-1(E)RI	2,432,410	2,141,052	88.02%	
D-15	1,170,000	121,997	10.43%	
D-20	9,375,000	7,397,508	78.91%	Actual workload includes D-20A 2nd Print Run
D-20(S)	892,500	137,583	15.42%	Actual workload includes D-20(S)A 2nd Print Run
D-21	500,000	288,753	57.75%	Actual workload includes D-21A 2nd Print Run
D-23	150,000	91,794	61.20%	Actual workload includes D-23A 2nd Print Run
D-1(XA)	91,196	163,730	179.54%	
D-1(XB)	21,046	19,436	92.35%	
D-1(X1)	14,030	12,189	86.88%	
D-1(X2)	21,046	19,383	92.10%	
D-1(X3)	21,046	19,377	92.07%	
D-1(X4)	21,046	19,477	92.54%	
D-1(X5)	21,046	19,578	93.02%	
D-1(X6)	21,046	19,430	92.32%	
D-1(X7)	21,046	19,489	92.60%	
D-1(X8)	21,046	19,413	92.24%	
D-1(X9)	21,046	19,430	92.32%	
D-1(X10)	21,046	19,482	92.57%	
D-1(X11)	21,046	19,283	91.62%	
D-1(X12)	21,046	19,416	92.26%	

Form[7]	Estimated Capture Workload	Total Checked in to Data Capture through 9/7/2010	Percent of Expected Workload Checked in	NOTES
D-1(X13)	21,046	21,646	102.85%	
D-1(X14)	21,046	19,387	92.12%	
D-1(X15)	21,046	19,397	92.16%	
D-1(X16)	21,046	19,386	92.11%	
D-1(X17)	21,046	19,513	92.72%	
D-1(E)X1*	600,000	517,381	86.23%	*Workload increased by CR
D-1(E)X2*	600,000	511,626	85.27%	*Workload increased by CR
D-20(X1)*	120,000	99,911	83.26%	*Workload increased by CR
D-1 PR	6,000	338	5.63%	
D-1 PR (S)*	932,994	866,389	92.86%	*Error in workload of J.27. Original estimate was 932,994, J.27 erroneously showed 6,000
D-1(UL) PR (S)	52,650	63,521	120.65%	
D-1(E)PR(S)	682,784	1,031,102	151.01%	
D-1(E) (SUPP) PR (S)	34,000	18,690	54.97%	
D-1(E)(RI)PR(S)	15,000	25,587	170.58%	
D-15 PR(S)	18,000	143	0.79%	
D-20 PR(S)	62,500	33,487	53.58%	
D-21 PR	775		0.00%	
D-23 PR	500	1	0.20%	
D-10 PR	1,392	1,213	87.14%	
D-10 PR (S)	17,152	9,284	54.13%	
D-20 PR	46,000	2,741	5.96%	

Data source: Cost and progress, Cumulative Data Capture Progress Report, and Section J.27 of the DRIS contract

Table A2: GQV Workload Comparison

Form[8]	Estimated Capture Workload	Total checked in to Data Capture through 12/31/2009	Percent of Expected Workload in DRIS	NOTES
D-351(GQV)	792,000	901,859	113.87%	
D-351 CF (GQV)	19,800	193	0.97%	
D-351 HU (GQV)	247,500	12,002	4.85%	
D-351 NSL(GQV)	990,000	1,141,175	115.27%	Includes 2nd Print Run of NSL-A forms
D-351(GQV) PR (S)	8,000	5,768	72.10%	
D-351 CF (GQV) PR (S)	200	0	0.00%	
D-351 HU (GQV) PR (S)	2,500	12	0.48%	
D-351 NSL(GQV) PR (S)	10,000	6,404	64.04%	
TOTAL for GQV	**2,070,000**	**2,067,413**	**99.88%**	

Data source: Cost and progress, Cumulative Data Capture Progress Report, and Section J.27 of the DRIS contract

Table A3: Dress Rehearsal Forms Processed at NPC using the DRIS System

Form Type	Total cumulative Check-in
DX-1 (English)	245,548
DX-1(E/S) Bilingual	10,416
TOTAL	**255,874**

Data Source: Cost and Progress, 2008 Dress Rehearsal Mailback Return Check-In Questionnaire Check-In (not unduplicated) report

[8] See Appendix B for explanation of Form Types.

Appendix B: Form type Descriptions by Form Number

Form Number	Description
D-1	Mailback - (Initial, Replacement, Fulfillment, Update/Leave (addressed)) -English
D-1(UL)	Update Leave ADDs – English
D-1PR(S)	Update/Leave (initial, addressed) and fulfillment – Puerto Rico - Spanish
D-1PR	Fulfillment - Puerto Rico – English
D-1(UL)PR(S)	Update/Leave ADD - Puerto Rico - Spanish
D-1(C)	Fulfillment - Chinese (Simplified)
D-1(K)	Fulfillment – Korean
D-1(R)	Fulfillment – Russian
D-1(S)	Fulfillment – Spanish
D-1(V)	Fulfillment – Vietnamese
D-1(E/S)	Bilingual - (mailout/mail back and update/leave) - English/Spanish
D-10	Be Counted-English
D-10(C)	Be Counted-Chinese
D-10(K)	Be Counted-Korean
D-10(R)	Be Counted-Russian
D-10(S)	Be Counted-Spanish
D-10(V)	Be Counted-Vietnamese
D-10PR	Be Counted-Puerto Rico (English)
D-10PR(S)	Be Counted-Puerto Rico (Spanish)
D-1(E)	Enumerator – English
D-1(E)(SUPP)	Enumerator Continuation – English
D-1(E)RI	Enumerator Reinterview – English
D-1(E)PR(S)	Enumerator, Puerto Rico Spanish
D-1(E)(SUPP)PR(S)	Enumerator Continuation - Puerto Rico – Spanish
D-1(E)(RI)PR(S)	Enumerator Reinterview, Puerto Rico, Spanish
D-15	Enumeration at Transitory Locations - English
D-15PR(S)	Enumeration at Transitory Locations- Puerto Rico (Spanish)

Form Number	Description
D-20	Individual Census Report (ICR)-English
D-20(X1)	Experimental Individual Census Report – English
D-20(S)	Individual Census Report (ICR)-Spanish
D-20PR	Individual Census Report (ICR)- Puerto Rico (English)
D-20PR(S)	Individual Census Report (ICR)- Puerto Rico (Spanish)
D-21	Military Census Report (MCR)-English
D-21PR	Military Census Report (MCR)- Puerto Rico (English)
D-23	Shipboard Census Report (SCR)-English
D-23PR	Shipboard Census Report (SCR)- Puerto Rico (English)
D-351(GQV)	Group Quarters Validation (GQV) Questionnaire – English
D-351CF(GQV)	GQV -Correctional Facility Continuation – English
D-351HU(GQV)	GQV - Housing Unit Continuation – English
D-351NSL(GQV)	GQV - Non-survivor Label Form – English
D-351NSL-A(GQV)	GQV - Non-survivor Label Form – English (second print run)
D-351(GQV)PR(S)	GQV questionnaire for Puerto Rico – Spanish
D-351CF(GQV)PR(S)	GQV - Correctional Facility Continuation Puerto Rico – Spanish
D-351HU(GQV)PR(S)	GQV - Housing Unit Continuation Puerto Rico –Spanish
D-351NSL(GQV)PR(S)	GQV - Non Survivor Label Page Puerto Rico – Spanish
D-1(XA)	Experimental Control Panel 1, with Over count Coverage Questions – English
D-1(XB)	Experimental Control Panel 2, without Over count Coverage Question – English
D-1(X1)	Experimental Version 1 (Census 2000 Treatment), Used for both Initial & Replacement Mailing
D-1(X2)	Experimental Version 2 (Combined Race/Hispanic Origin Treatment), Used for both Initial and Replacement Mailing
D-1(X3)	Experimental Version 3 (Combined Race/Hispanic Origin Treatment), Used for both Initial and Replacement Mailing
D-1(X4)	Experimental Version 4 (Combined Race/Hispanic Origin Treatment), Used for both Initial and Replacement Mailing
D-1(X5)	Experimental Version 5 (Combined Race/Hispanic Origin Treatment), Used for both Initial and Replacement Mailing
D-1(X6)	Experimental Version 6 (Race Treatment), Used for both Initial and Replacement Mailing

Form Number	Description
D-1(X7)	Experimental Version 7 (Hispanic Origin Treatment), Used for both Initial and Replacement Mailing
D-1(X8)	Experimental Version 8 (Hispanic Origin Treatment), Used for both Initial and Replacement Mailing
D-1(X9)	Experimental Version 9 (Race and Hispanic Origin Treatment), Used for both Initial and Replacement Mailing
D-1(X10)	Experimental Version 10 (Race and Hispanic Origin Treatment), Used for both Initial and Replacement Mailing
D-1(X11)	Experimental Version 11 (Race and Hispanic Origin Treatment), Used for both Initial and Replacement Mailing
D-1(X12)	Experimental Version 12 (Hispanic Origin Treatment) Used for both Initial and Replacement Mailing
D-1(X13)	Experimental Version 13 (Coverage Treatment) Used for both Initial and Replacement Mailing (booklet)
D-1(X14)	Experimental Version 14 Office of Management and Budget Race Panel 1 - English
D-1(X15)	Experimental Version 15 – Office of Management and Budget Race Panel 2 – English
D-1(X16)	Experimental Version 16 - Office of Management and Budget Race Panel 3 – English
D-1(X17)	Experimental Version 17 - Office of Management and Budget Race Panel 4 – English
D-1(E)X1	Experimental Enumerator 5-contact -- English
D-1(E)X2	Experimental Enumerator 4-contact-- English
UNKNOWN	Form type could not be determined
Data Source: Cost and Progress Database Design Document	

APPENDIX C: Undeliverable As Addressed (UAA) Exception Check-In Rates Table

Form Type	Cumulative UAA Sorter Check-In	Cumulative UAA Exception Check-In	Total Number of UAAs Checked-In Cumulative	UAA Exception Check-In Rate
D-1	18,003,853	360,610	18,364,463	1.96%
D-1(C)	64	11	75	14.67%
D-1(E/S)	1,023,381	98,117	1,121,498	8.75%
D-1(K)	42	8	50	16.00%
D-1(R)	31	6	37	16.22%
D-1(S)	4,778	671	5,449	12.31%
D-1(UL)	1,802	263	2,065	12.74%
D-1(UL)PR(S)	383	81	464	17.46%
D-1(V)	36	1	37	2.70%
D-1PR	35	8	43	18.60%
D-1PR(S)	1,434	166	1,600	10.38%
D-1(X1)	2,725	34	2,759	1.23%
D-1(X10)	3,253	50	3,303	1.51%
D-1(X11)	3,418	61	3,479	1.75%
D-1(X12)	3,305	68	3,373	2.02%
D-1(X13)	3,207	500	3,707	13.49%
D-1(X14)	3,434	45	3,479	1.29%
D-1(X15)	3,412	53	3,465	1.53%
D-1(X16)	3,311	54	3,365	1.60%
D-1(X17)	3,278	64	3,342	1.92%
D-1(X2)	3,385	54	3,439	1.57%
D-1(X3)	3,521	52	3,573	1.46%
D-1(X4)	3,372	55	3,427	1.60%
D-1(X5)	3,373	65	3,438	1.89%
D-1(X6)	3,317	61	3,378	1.81%
D-1(X7)	3,376	68	3,444	1.97%
D-1(X8)	3,292	60	3,352	1.79%
D-1(X9)	3,309	62	3,371	1.84%
D-1(XA)	37,030	520	37,550	1.38%
D-1(XB)	3,311	59	3,370	1.75%
TOTAL	**19,132,468**	**461,927**	**19,594,395**	**2.36%**

Data source: Cost and Progress, UAA Check-in status by Data Capture Center Report

Appendix D: Exception Check-In Rates for Mail Returns by Form Type

Form Type	Total Form Check-In Cumulative	# Forms Exception Check-In Cumulative	Exception Check-in Rate
D-1	82,929,578	2,501,072	3.02%
D-1(C)	1,535	219	14.27%
D-1(E/S)	7,273,757	216,739	2.98%
D-1(K)	1,417	235	16.58%
D-1(R)	558	75	13.44%
D-1(S)	55,933	6,875	12.29%
D-1(UL)	271,535	13,781	5.08%
D-1(UL)PR(S)	63,521	4,819	7.59%
D-1(V)	1,180	165	13.98%
D-10	681,320	37,005	5.43%
D-10(C)	10,871	525	4.83%
D-10(K)	4,680	255	5.45%
D-10(R)	2,071	218	10.53%
D-10(S)	71,246	4,916	6.90%
D-10(V)	3,335	239	7.17%
D-10PR	1,213	92	7.58%
D-10PR(S)	9,284	693	7.46%
D-1PR	338	42	12.43%
D-1PR(S)	866,389	76,025	8.77%
D-1X1	12,189	600	4.92%
D-1X10	19,482	757	3.89%
D-1X11	19,283	739	3.83%
D-1X12	19,416	763	3.93%
D-1X13	21,646	1,260	5.82%
D-1X14	19,387	740	3.82%
D-1X15	19,397	752	3.88%
D-1X16	19,386	799	4.12%
D-1X17	19,513	767	3.93%
D-1X2	19,383	777	4.01%
D-1X3	19,377	781	4.03%
D-1X4	19,477	790	4.06%
D-1X5	19,578	771	3.94%
D-1X6	19,430	743	3.82%
D-1X7	19,489	781	4.01%
D-1X8	19,413	748	3.85%
D-1X9	19,430	787	4.05%
D-1XA	163,730	6,052	3.70%
D-1XB	19,436	760	3.91%
TOTAL	**92,758,203**	**2,884,157**	**3.11%**

Form Type	Total Form Check-In Cumulative	# Forms Exception Check-In Cumulative	Exception Check-in Rate
Data Source:	Cost and Progress, Cumulative Data Capture Progress Report		

APPENDIX E: List of System Data Exchanges between DRIS and Census

#	Interface Partner	System Data Exchange ID	Interface Title
1	Field Data Collection Automation (FDCA)	DRIS:FDCA-1	LCO Box Receipt Confirmation
2	FDCA	DRIS:FDCA-2	Universe Transaction Update
3	FDCA	DRIS:FDCA-3	GQ Form Receipt Confirmation
4	FDCA	DRIS:FDCA-4	Form Receipt Confirmation – for GQV Operation
5	FDCA	DRIS:FDCA-4	Form Receipt Confirmation for not GQV or GQE
6	FDCA	FDCA:DRIS-1	LCO Box Shipment Notification
7	FDCA	FDCA:DRIS-3	GQ Form Identification Notification
8	FDCA	FDCA:DRIS-4	GQV Universe
9	FDCA	FDCA:DRIS-5	GQ ADD to AA Linkage
10	FDCA	FDCA:DRIS-6	Parent to Continuation Form Linkage Notification
11	RR Donnelley (RRD)	DRIS:RRD-1	Replacement Mailing Data Delivery
12	RRD	DRIS:RRD-2	CPEX Replacement Mailing Data Delivery
13	RRD	RRD:DRIS-1	Origin Confirmation (Pre-Check-in)
14	RRD	RRD:DRIS-2	Destination Confirmation (UAAs)
15	RRD	RRD:DRIS-3	Mail sort Rejects (UAAs)
16	RRD	RRD:DRIS-4	Sectional Center

#	Interface Partner	System Data Exchange ID	Interface Title
			Facility List
17	Coverage Followup (CFU)	DRIS:CFU-1	CFU Metrics Reports
18	CFU	DRIS:CFU-2	CFU Auxiliary Data
19	CFU	DRIS:CFU-3	Telephony Contact History File (CFU Call Detail Record)
20	CFU	DRIS:CFU-4	Module A/P Data (Household/Person Data)
21	CFU	DRIS:CFU-5	CFU Recordings
22	Decennial Statistical Studies Division: Evaluations, Experiments & Assessments-- Evaluations (DSSD: EE&A – Assessments)	DRIS:CEE-16	Phone Numbers for AQE/Content Reinterview Cases
23	DSSD: EE&A – Assessments	DRIS:CEE-17	TQA Data Before and After Data Review (Pre & Post)
24	DSSD: EE&A – Assessments	DSSD:DRIS-1	AQE/Content Reinterview Case IDs
25	Headquarters Processing (HQP)	DRIS:RPS-1	Response Delivery
26	HQP	DRIS:UC&M-1	Universe Status Delivery
27	HQP	RPS:DRIS-1	Response Data Confirmation
28	HQP	UC&M:DRIS-1	Universe
29	HQP	UC&M:DRIS-2	Universe Status Exception List
30	HQP	UC&M:DRIS-3	Universe Mailing Files
31	Production Data Quality (PDQ)	DRIS:PDQ-1	PDQ Data
32	PDQ	PDQ:DRIS-1	PDQ Metrics
33	PDQ	PDQ:DRIS-2	PDQ History
34	Cost and Progress (C&P)	DRIS:C&P-1	Cost and Progress Data
35	C&P	C&P:DRIS-1	VDC Contingency Linkage Files
36	Person MaRCS (Per	DRIS: Per MaRCS-1	CCM Images

#	Interface Partner	System Data Exchange ID	Interface Title
	MaRCS)		
37	National Processing Center (NPC)	DRIS:NPC-1	Form Receipt Confirmation
38	NPC	NPC:DRIS-1	GQ to Form Linkage File
39	PBO	DRIS:PBO-1	LCO Box Receipt Confirmation
40	PBO	DRIS:PBO-2	Universal Transaction Update
41	PBO	DRIS:PBO-4	Form Receipt Confirmation
42	PBO	PBO:DRIS-1	LCO Box Shipment Notification
43	PBO	PBO:DRIS-6	Parent to Child Form Linkage Notification
44	Census MaRCS	DRIS: Census MaRCS-1	NRFU and UE Response and Auxiliary Data
45	Census MaRCS	DRIS: Census MaRCS-2	Note Images
46	Census MaRCS	Census MaRCS:DRIS-1	NRFU and UE Response Data Confirmation
47	DSSD: EE&A – Assessments	DRIS:CEE-10	Assessment Data
48	DSSD: EE&A – Assessments	DRIS:CEE-13	Call Detail Data
49	DSSD: EE&A – Assessments	DRIS:CEE-11	SQA Metrics
50	DSSD: EE&A – Experiments	DRIS:CEE-6	Responses to CPEX Experimental Forms
51	DSSD: EE&A – Experiments	DRIS:CEE-7	TQA Line CPEX Callers
52	DSSD: EE&A – Experiments	DRIS:CEE-14	CPEX Replacement Mailing File
53	DSSD: EE&A – Experiments	DRIS:CEE-15	Raw CPEX Form Response Data
54	DSSD Data Quality (DQ)	DRIS:CEE-1a	Data Capture QA Field Data
55	DSSD Data Quality (DQ)	DRIS:CEE-1b	Data Capture QA Batch Data
56	DSSD Data Quality	DRIS:CEE-2	CFU DQ Progress

#	Interface Partner	System Data Exchange ID	Interface Title
	(DQ)		Reports
57	DSSD Data Quality (DQ)	DRIS:CEE-3	CFU DQ Data Dump
58	Census Program for Evaluations & Experiments	DRIS:CEE-4	WCM System metrics report
59	DSSD Data Quality (DQ)	DRIS:CEE-18	Original PDQ Data
60	DSSD Data Quality (DQ)	DRIS:CEE-19	PDQ History Pass-through
61	DSSD: Other	DRIS:CEE-20	Check-In Data
62	DSSD: Other	DRIS:CEE-21	Late Response Data
63	DSSD: Census Bureau	DRIS:CB-1	Census Questionnaire Images Delivery
64	Census Image Retrieval Application (CIRA)	DRIS:CIRA-1	Encryption Key, Box ID number, FedEx Tracking Number
65	CIRA	DRIS:CIRA-2	Inventory File
66	CIRA	DRIS:CIRA-3	Images
67	CIRA	DRIS:CIRA-4	Manifest
68	CIRA	DRIS:CIRA-5	Control File
69	CIRA	CIRA:DRIS-1	Encryption Key Receipt
70	CIRA	CIRA:DRIS-2	Box Receipt Confirmation
71	CIRA	CIRA:DRIS-3	Resend File
72	CIRA	CIRA:DRIS-4	Cumulative Tape Processing Report
73	CIRA	CIRA:DRIS-5	Daily Tape Status

Source: DRIS Program Management Office