

January 2005

2010 CENSUS

Basic Design Has Potential, but Remaining Challenges Need Prompt Resolution



G A O

Accountability * Integrity * Reliability

GAO
Accountability · Integrity · Reliability

Highlights

Highlights of [GAO-05-9](#), a report to congressional requesters

Why GAO Did This Study

A rigorous testing and evaluation program is a critical component of the census planning process because it helps the U.S. Census Bureau (Bureau) assess activities that show promise for a more cost-effective head count. The Bureau conducted a field test in 2004, and we were asked to (1) assess the soundness of the test design and the extent to which the Bureau implemented it consistent with its plans, (2) review the quality of the Bureau's information technology (IT) security practices, and (3) identify initial lessons learned from conducting the test and their implications for future tests and the 2010 Census.

What GAO Recommends

We recommend that the Secretary of Commerce direct the Bureau to address the shortcomings revealed during the 2004 test. Specific actions include enhancing the Bureau's IT security practices; improving the reliability of hand-held computer (HHC) transmissions; developing a more strategic approach to training; and ensuring that all systems are test ready. The Bureau should also regularly update Congress on its progress in addressing these issues and meeting its 2010 goals. The Bureau generally agreed with most of our recommendations, but took exception to two of them concerning certain census activities and their impact on Bureau objectives. However, we believe those recommendations still apply.

www.gao.gov/cgi-bin/getrpt?GAO-05-9.

To view the full product, including the scope and methodology, click on the link above. For more information, contact Patricia A. Dalton at (202) 512-6806 or daltonp@gao.gov.

2010 CENSUS

Basic Design Has Potential, but Remaining Challenges Need Prompt Resolution

What GAO Found

The Bureau's design for the 2004 census test addressed important components of a sound study, and the Bureau generally implemented the test as planned. For example, the Bureau clearly identified its research objectives, developed research questions that supported those objectives, and developed evaluation plans for each of the test's 11 research questions.

The initial results of the test suggest that while certain new procedures show promise for improving the cost-effectiveness of the census, the Bureau will have to first address a number of problems that could jeopardize a successful head count. For example, enumerators had little trouble using hand held computers (HHC) to collect household data and remove late mail returns. The computers could reduce the Bureau's reliance on paper questionnaires and maps and thus save money. The test results also suggest that certain refinements the Bureau made to its procedures for counting dormitories, nursing homes, and other "group quarters" could help prevent the miscounting of this population group.

The 2004 Census Test Was Conducted in Rural Georgia and Queens, New York

COLQUITT, THOMAS, and TIFT COUNTIES...

Census Workers are in Your Neighborhood

If the U.S. Census Bureau did not receive your 2004 Census Test form, a census worker will come to your door. Please answer their questions.

Census Worker's Tools:

Badge, Handheld Computer, Bag

2004 CENSUS TEST

For more information about the 2004 Census Test, call 877-648-9819

USCENSUSBUREAU

ASTORIA, CORONA, ELMHURST, EAST ELMHURST, JACKSON HEIGHTS, LONG ISLAND CITY, WOODSIDE

Census Workers are in Your Neighborhood

If the U.S. Census Bureau did not receive your 2004 Census Test form, a census worker will come to your door. Please answer their questions.

Census Worker's Tools:

Badge, Handheld Computer, Bag

2004 CENSUS TEST

For more information about the 2004 Census Test, call 877-648-9819

USCENSUSBUREAU

Source: U.S. Census Bureau.

The Bureau conducted a 2004 census test in Queens, New York and three counties in rural Georgia. The test focused on using HHCs for conducting nonresponse follow-up interviews—where census workers go door-to-door to count those households that did not mail back their census forms. The Bureau also tested new methods for improving coverage, redesigned race and ethnicity questions, and improved methods for defining and identifying group quarters. The Bureau established these test objectives to help meet its goals for improving accuracy, reducing risk, and containing the cost of the 2010 Census.

Other aspects of the test did not go as smoothly. For example, security practices for the Bureau's IT systems had weaknesses; the HHCs had problems transmitting data; questionnaire items designed to improve coverage and better capture race/ethnicity confused respondents; enumerators sometimes deviated from prescribed enumeration procedures; and certain features of the test were not fully operational at the time of the test, which hampered the Bureau from fully gauging their performance. With few testing opportunities remaining, it will be important for (1) the Bureau to find the source of these problems, devise cost-effective solutions, and integrate refinements before the next field test scheduled for 2006, and (2) Congress to monitor the Bureau's progress in resolving these issues.

Contents

Letter

Results in Brief	1
Background	2
Scope and Methodology	4
The Census Test Was Generally Sound, but Refinements Could Produce Better Cost and Performance Data	6
The Bureau Needs to Implement Better IT Security Practices	8
Test Reveals Technical, Training, and Other Challenges in Need of Prompt Resolution	11
Conclusions	13
Recommendations for Executive Action	29
Agency Comments and Our Evaluation	30
	31

Appendix

Appendix I: Comments from the Department of Commerce	34
---	-----------

Table

Table 1: Design for 2004 Census Test Addressed Important Components of a Sound Study	9
--	---

Figures

Figure 1: HHC Being Tested for Use in Collecting Data in the Field	5
Figure 2: Maps of Test Sites in Georgia and New York	7
Figure 3: An Enumerator Using an HHC for Nonresponse Follow-up	14
Figure 4: Data Transmission Process for Nonresponse Follow-up	16
Figure 5: New Coverage Questions Were Designed to Ensure a Complete Count	20
Figure 6: Race and Ethnicity Categories on the HHCs Were Formatted Differently From the Paper Questionnaires	22
Figure 7: Group Homes Could Resemble Conventional Houses	25

This is a work of the U.S. government and is not subject to copyright protection in the United States. It may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.



United States Government Accountability Office
Washington, D.C. 20548

January 12, 2005

The Honorable Tom Davis
Chairman
The Honorable Henry A. Waxman
Ranking Minority Member
Committee on Government Reform
House of Representatives

The Honorable Adam H. Putnam
House of Representatives

The consequences of a poorly planned census are high given the billions of dollars spent on the enumeration and the importance of collecting quality data. Therefore, a rigorous testing and evaluation program is a critical component of the census planning process because it helps the U.S. Census Bureau (Bureau) assess activities and related information technology (IT) systems that show promise for a more cost-effective head count. In preparing for the 2010 Census, the Bureau conducted a field test in 2004 and plans additional tests for 2005 and 2006. It is important that these early assessments lead to a design that is sufficiently mature so that the dress rehearsal for the 2010 Census, now planned for 2008, will demonstrate the feasibility of the various operations and technologies planned for the decennial under conditions that are as close as possible to the actual census.

The Bureau designed the 2004 census test to examine the feasibility of using (1) handheld computers (HHC) for field data collection; (2) new methods for improving coverage; (3) redesigned race and ethnicity (Hispanic origin) questions; and (4) improved methods for defining and identifying nursing homes, prisons, college dormitories, and similar facilities known collectively as “group quarters.” The Bureau established these test objectives as part of a broader effort to modernize and re-engineer the 2010 Census. Major goals of this initiative are to improve the accuracy, reduce the risk, and contain the cost of the 2010 Census, estimates of which now exceed \$11 billion. A rigorous planning and testing program is critical to this effort.

As agreed with your offices, our objectives for this review were to assess the soundness of the design of the 2004 census test and the extent to which the Bureau implemented the test consistent with its plans. We also agreed to review the quality of its IT security practices, and identify initial lessons

learned from the test and the implications they have for the Bureau's future plans.

To address these objectives, we reviewed applicable planning, IT, and other documents, and interviewed knowledgeable Bureau officials responsible for key operations and computer security. We also made several visits to the two test sites—an urban location in the northwestern portion of Queens Borough, New York, and a rural location in south central Georgia. We conducted our work from November 2003 through November 2004 in accordance with generally accepted government auditing standards.

Results in Brief

The design of the 2004 census test addressed important components of a sound study, and the Bureau generally implemented the test as planned. For example, the Bureau clearly identified its test objectives, developed research questions that supported those objectives, and developed evaluation plans for each of the test's 11 research questions.

Still, there are opportunities to improve both the utility of data from the current test as well as the design of the next field test in 2006. Combined, these improvements could help inform future budget estimates and investment and design decisions. For example, the 2004 test could benefit by analyzing the impact the HHCs and targeted second mailing—an operation designed to increase the response rate—had on cost savings and productivity. Similarly, the 2006 test could be improved if the Bureau developed quantifiable productivity and other performance requirements for the HHCs and then used the 2006 test to determine whether the devices are capable of meeting those requirements.

The 2004 test was an important milestone in the 2010 Census life cycle because it shed light on those operations that show potential for improving the cost-effectiveness of the decennial head count, as well as problem areas that could jeopardize the success of the census if not resolved. For example, the initial test results showed that the HHCs were effective for conducting interviews and removing late mail returns. Indeed, most enumerators we observed had little trouble using the computers for conducting interviews, and they were generally pleased with the HHC's overall functionality, durability, and screen clarity. Likewise, the HHCs enabled the Bureau to remove over 7,000 late mail returns from enumerators' workloads at both test sites, which could help the Bureau save money by eliminating the need to visit those households that already mailed back their census questionnaires. The test results also suggest that

certain refinements the Bureau made to its procedures for counting group quarters—namely integrating its housing unit and group quarters address lists—could help prevent the miscounting of this population group.

Other aspects of the test did not go as smoothly and point to areas on which the Bureau should focus as it looks toward the future. For example:

- The Bureau's IT security practices had weaknesses;
- Technical and training difficulties caused HHC transmission problems;
- The HHC's mapping function was slow to load and was thus little used;
- Questionnaire items designed to improve coverage and better determine race/ethnicity were awkward for census workers to ask and confusing for respondents, which could affect data quality;
- Census workers sometimes deviated from prescribed enumeration procedures, which could impair the reliability of the data;
- Enumerators had difficulties finding the physical locations of specific group quarters; and
- Certain features of the test were not fully operational at the time of the test, which hampered the Bureau from gauging their true performance.

The 2006 field test is the Bureau's last opportunity to assess its basic design for the census before conducting a dress rehearsal in 2008. At that time, the Bureau plans to demonstrate the entire design under conditions that mirror the census. Any changes to the design made after the dress rehearsal could entail considerable risk as they would not be properly tested. Thus, it will be important for the Bureau to exhaustively assess the results of the 2004 test, diagnose and remedy any shortcomings, and thoroughly road test refinements in 2006.

To facilitate this, we recommend that the Secretary of Commerce direct the Bureau to address the technical, methodological, training, and procedural shortcomings revealed during the 2004 test. Specific actions include enhancing the Bureau's IT security practices, improving the reliability of HHC transmissions, and taking a more strategic approach to training enumerators. The Bureau should also regularly update Congress on the progress it is making in addressing these and any other challenges, as well

as the extent to which it is on track for meeting the overall goals of the 2010 Census.

The Under Secretary for Economic Affairs forwarded written comments from the Bureau on a draft of this report (see app. I). The Bureau generally agreed with seven of our nine recommendations—those dealing with improving IT security practices, the reliability of the HHCs, training, testing, and enumeration procedures—and reported it was already taking a number of steps to address our concerns. However, the Bureau took exception to our recommendations to (1) analyze the impact that HHCs and a targeted second mailing had on cost savings and other Bureau objectives, and (2) define specific, measurable performance requirements for the HHCs and other census-taking activities. Because the HHCs and certain other operations are critical for containing costs and achieving other Bureau goals for the 2010 Census, it will be essential for the Bureau to gauge their impact and determine whether they can meet the Bureau’s demanding requirements. As a result, we believe these recommendations still apply.

Background

Congress, GAO, the Department of Commerce Inspector General, and even the Bureau itself have all noted that the 2000 Census was marked by poor planning, which unnecessarily added to the cost, risk, and controversy of the national head count. In January 2003, we named the 2010 Census a major performance and accountability challenge because of our growing concern over the numerous obstacles to a cost-effective enumeration as well as its escalating price tag.¹ More recently, we reported that while the Bureau’s preparations for the 2010 Census appeared to be further along than at a similar point during the planning cycle for the 2000 Census, considerable risks and uncertainties remained.² Thus, it is imperative that the Bureau adequately test the various components of its design for the 2010 Census.

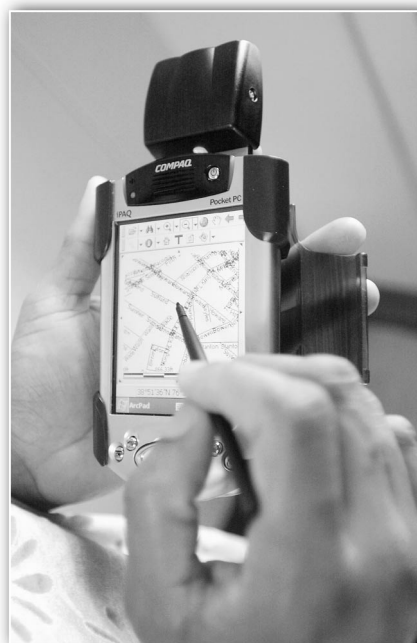
A rigorous testing program provides at least four major benefits. First, testing allows the Bureau to refine procedures aimed at addressing

¹ GAO, *Major Management Challenges and Program Risks: Department of Commerce*, [GAO-03-97](#) (Washington, D.C.: January 2003).

² GAO, *2010 Census: Cost and Design Issues Need to Be Addressed Soon*, [GAO-04-37](#) (Washington, D.C.: Jan. 15, 2004).

problems encountered in past censuses. During the 2000 Census, for example, group quarters were sometimes counted more than once or counted in the wrong location; the wording of the race and ethnicity question confused some respondents, which in some cases resulted in lower quality data; and following up with nonrespondents proved to be costly and labor-intensive. A second benefit is that sound testing can assess the feasibility of new procedures and technologies, such as HHCs (see fig. 1), that have never before been used in a decennial census.

Figure 1: HHC Being Tested for Use in Collecting Data in the Field



Source: U.S. Census Bureau.

Third, a rigorous testing program helps instill a comfort level among members of Congress and other stakeholders that the Bureau (1) has chosen the optimal design given various trade-offs and constraints and (2) has identified and addressed potential risks and will be able to successfully execute its plan. Such confidence building, developed through regular updates and open lines of communication, is essential for continuing congressional support and funding.

And finally, proper testing early in the decade will help the Bureau to conduct a dress rehearsal in 2008 that fully assesses all aspects of the census design under realistic conditions. Because of various late requirement changes, certain procedures that were added after the 1998 dress rehearsal for the 2000 Census were not properly tested.

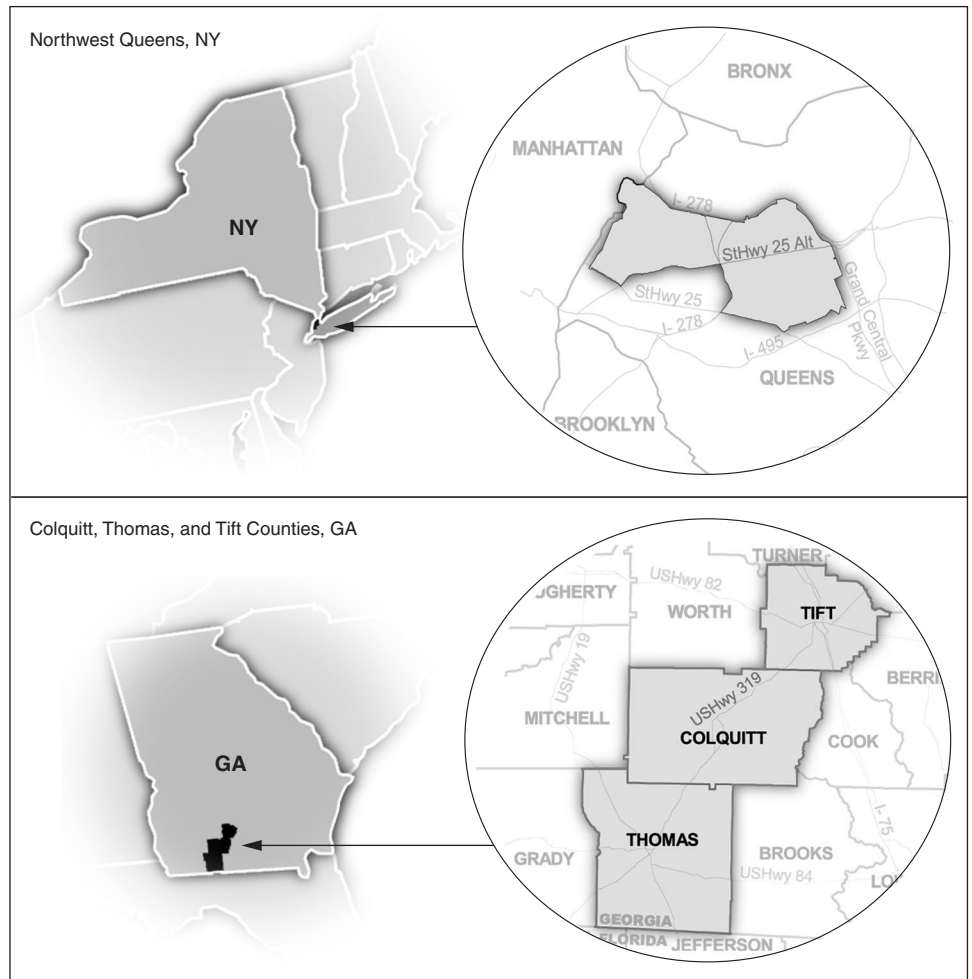
Scope and Methodology

As agreed with your offices, our objectives for this report were to (1) assess the soundness of the Bureau's design for the 2004 census test and whether the Bureau implemented the test consistent with its plans, (2) review the quality of the Bureau's IT security practices, and (3) identify initial lessons learned from conducting the test and their implications for the 2010 Census.

To assess the soundness of the design we reviewed pertinent documents that described the Bureau's test and evaluation plans. We systematically rated the Bureau's approach using a checklist of design elements that, based on our review of program evaluation literature, are relevant to a sound study plan. For example, we reviewed the Bureau's approach to determine, among other things, (1) how clearly the Bureau presented research objectives, (2) whether research questions matched the research objectives, and (3) the appropriateness of the data collection strategy for reaching the intended sample population. As part of our assessment of the Bureau's test design, we also reviewed evaluations of the prior decennial census to determine the degree to which the new operations being tested addressed problematic aspects of the 2000 Census. However, we did not assess the Bureau's criteria in selecting its objectives for the 2004 census test.

To determine if the Bureau implemented the test consistent with its plans, we made multiple site visits to local census offices in Thomasville, Georgia; and Queens Borough, New York. During these visits, we interviewed local census office managers and staff, observed various data collection activities, and attended weeklong enumerator training. We observed a total of 20 enumerators as they completed their daily nonresponse follow-up assignments—half of these were in southern Georgia, in the counties of Thomas, Colquitt, and Tift, and half were in Queens (see fig. 2 for maps of the test site areas). The results of these observations are not necessarily representative of the larger universe of enumerators.

Figure 2: Maps of Test Sites in Georgia and New York



Source: GAO and U.S. Census Bureau.

To evaluate the quality of the Bureau's IT security practices, we assessed risk management documentation associated with IT systems and major applications for the 2004 census test. We based our determination on applicable legal requirements, Bureau policy, and leading practices described in our executive guide for information security management.³

³GAO, *Executive Guide: Information Security Management—Learning from Leading Organizations*, [GAO/AIMD-98-68](#) (Washington, D.C.: May 1998).

We also interviewed key Bureau officials associated with computer security.

To identify lessons learned from the 2004 census test, we met with officials from the Bureau's Decennial Management Division regarding overall test plans and with officials from its Technologies Management Office about using HHCs. Bureau officials and census workers from both test locations also provided suggestions on improving census operations.

We requested comments on a draft of this report from the Secretary of Commerce. On December 20, 2004, the Under Secretary for Economic Affairs, Department of Commerce, forwarded written comments from the Bureau (see app. I). We address these comments in the "Agency Comments and Our Evaluation" section at the end of this report.

The Census Test Was Generally Sound, but Refinements Could Produce Better Cost and Performance Data

The Bureau designed a sound census test and generally implemented it as planned. However, in looking ahead, the Bureau's planning and investment decisions could benefit from analyzing (1) the degree to which HHCs contributed to the Bureau's cost containment goal and (2) the results of the targeted second mailing, an operation designed to increase participation by sending a follow-up questionnaire to nonresponding households. Future tests could also be more informative if the Bureau developed quantifiable productivity and other performance requirements for the HHCs and then used the 2006 test to determine whether the devices are capable of meeting those requirements. Collectively, these refinements could provide Bureau officials with better information to guide its IT and other design decisions, as well as refine future census tests.

The Bureau Developed a Sound Test Design

The design of the 2004 census test contained many components of a sound study (see table 1). For example, the Bureau identified test objectives, designed related research questions, and described a data collection strategy appropriate for a field test. The Bureau also developed evaluation plans for each of the test's 11 research questions, and explained how stakeholders were involved with the design, as well as how lessons learned from past studies were incorporated.

Table 1: Design for 2004 Census Test Addressed Important Components of a Sound Study

Components of a sound study	Planned components of the 2004 census test
Clearly stated objectives	The objectives for the test concerned the feasibility of using: <ul style="list-style-type: none"> • HHCs for field data collection; • new methods for improving coverage; • redesigned race and ethnicity questions; and • improved methods for defining and identifying group quarters.
Research questions linked to objectives and rationale for site selection provided	<ul style="list-style-type: none"> • Each of the 11 key research questions can be linked to one of the four objectives. • Two sites—in Queens and south central Georgia—were selected based on test requirements.
Data collection strategy thoroughly documented	<ul style="list-style-type: none"> • Census Day was April 1, 2004, for the test. • Mode of data collection was the paper version of the short form to be mailed back by household. • Nonrespondent data were collected during personal interviews using HHCs.
Input from stakeholders and lessons learned considered in developing test objectives	<ul style="list-style-type: none"> • Various research and development planning groups were convened to develop 2004 test objectives. • The Bureau’s Decennial Advisory Board and the National Academy of Science were informed of test plans. • Lessons learned from the 2000 Census were considered in developing 2004 test objectives.
Design had data analysis plan	<ul style="list-style-type: none"> • Separate evaluations to answer key research questions were developed. • Evaluation plans recognized limitations. For example, the introduction of a new data collection mode and new questions may make comparison to 2000 data difficult.

Source: GAO analysis of U.S. Census Bureau data.

Additional Analysis Would Provide Better Data on the Impact of Key Test Components

Although the Bureau plans to evaluate various aspects of the 2004 test, it does not currently plan to assess the impact that the HHCs and targeted second mailing had on cost savings and productivity. According to the Bureau, the census test was focused more on determining the feasibility of using the HHCs and less on the devices’ ability to save money. Likewise, the Bureau said it is not assessing the impact of the targeted second mailing because the operation is not one of its four test objectives for improving (1) field data collection using the HHC, (2) the coverage of undercounted groups, (3) questions about race and ethnicity, and (4) methods for defining special places and group quarters.

These decisions might be shortsighted, however, in that the Bureau included the HHCs and targeted second mailing in the 2010 Census design, in part, to reduce staff, improve productivity, and control costs. For example, Bureau studies have shown that sending out replacement questionnaires could yield a gain in overall response of 7 to 10 percent from households that do not respond to the initial census mailing, and thus generate significant cost savings by eliminating the need for census workers to obtain those responses via personal visits. Thus, information on the degree to which the HHCs and second mailing contribute to these key goals could help inform future budget estimates, investment and design decisions, as well as help refine future census tests.

Moreover, the feasibility of a targeted second mailing is an open question, as the Bureau has never before included this operation as part of a decennial census. Although a second mailing was part of the original design for the 2000 Census, the Bureau had abandoned it because it was found to be logistically unworkable. A Bureau official said that the second mailing was included in the 2004 test only to facilitate the enumeration process, and it would be better tested in a larger scale operation such as the 2008 dress rehearsal. However, we believe that it would be more prudent to assess the second mailing earlier in the census cycle, such as, during the 2006 test so that its basic feasibility could be assessed, any refinements could be evaluated in subsequent tests, and the impact on savings could be estimated more accurately.

Future Tests Could Be Improved

While the design of the 2004 test was generally sound, refinements could strengthen the next field test in 2006. Opportunities for improvement exist in at least two areas: ensuring that (1) the HHCs can meet the demanding requirements of field data collection and (2) management of the local census offices mirrors an actual enumeration as much as possible.

With respect to the HHCs, because they replace the paper version of the nonresponse follow-up questionnaire the devices must function effectively. Further, this test was the first time the Bureau used the HHCs under census-like conditions so their functionality in an operational environment was unknown. Bureau officials have acknowledged that for the 2004 test they had no predefined indicators of success or failure other than if there was a complete breakdown the test would be halted. This is a very low standard. Now that the Bureau has demonstrated the basic functionality of the computers, it should next focus on determining the specific performance requirements for the HHCs and assess whether the devices

are capable of meeting them. For example, the Bureau needs productivity benchmarks for the number of interviews per hour and per day that is expected per census worker. Durability measures, such as how many devices were repaired or replaced, should be considered as well. Assessing whether the HHCs can meet the requirements of nonresponse follow-up will help inform future design and investment decisions for whether or not to include the devices in the 2010 design.

Ensuring that key positions in the local census offices are filled from the same labor pool as they would be in an actual decennial census could also enhance future census tests. Such was not the case during the 2004 test when, according to the Bureau, because of difficulties finding qualified applicants, it used an experienced career census employee to manage the overall day-to-day operations of the local census office at the Queens test site. Another career employee occupied the office's regional technician slot, whose responsibilities included providing technical and administrative guidance to the local census office manager. In the actual census, the Bureau would fill these and other positions with temporary employees recruited from local labor markets. However, because the Bureau staffed these positions with individuals already familiar with census operations and who had ties to personnel at the Bureau's headquarters, the Queens test may not have been realistic and the test results could be somewhat skewed.

The Bureau Needs to Implement Better IT Security Practices

The Bureau operated a number of IT systems in order to transmit, manage, and process data for the test. The equipment was located at various places including the Bureau's headquarters in Suitland, Maryland; its National Processing Center in Jeffersonville, Indiana; a computer facility in Bowie, Maryland; as well as the New York and Georgia test sites.

Under Title 13 of the U.S. Code, the Bureau must protect from disclosure the data it collects about individuals and establishments. Thus, the Bureau's IT network must support both the test's telecommunications and data processing requirements, as well as safeguard the confidentiality and integrity of respondents' information.

The Federal Information Security Management Act of 2002 (FISMA) requires each agency to develop, document, and implement an agency-wide information security program for the IT systems that supports its operations.⁴ Although the Bureau took a number of steps to implement IT security over the systems used for the test, based on available information, the Bureau did not meet several of FISMA's key requirements. As a result, the Bureau could not ensure that the systems supporting the test were properly protected against intrusion or unauthorized disclosure of sensitive information. For example:

- **IT inventory was not complete.** FISMA requires an inventory of major information systems and interfaces. The Bureau did not have a complete inventory that showed all applications and general support IT systems associated with the test. Without such information, the Bureau could not ensure that security was effectively implemented for all of its systems used in the test, including proper risk assessments, adequate security plans, and effectively designed security controls.
- **There was not sufficient evidence that the Bureau assessed all of the devices used in the test for vulnerabilities, or that it corrected previously identified problems.** FISMA requires that agencies test and evaluate the effectiveness of information security policies, procedures, and practices for each system at least annually and that agencies have a process for remediating any identified security weaknesses. Since the Bureau could not provide us with a complete inventory of all network components used in the test, we could not determine if the Bureau's tests and evaluations were complete. Moreover, there was not always evidence about whether the Bureau had corrected past problems or documented reasons for not correcting them. As a result, the Bureau did not have adequate assurance that the security of systems used in the 2004 census test was adequately tested and evaluated or that identified weaknesses were corrected on a timely basis.
- **Assessments were not consistent.** FISMA requires agencies to assess the risks that could result from the unauthorized access, use, disclosure, disruption, modification, or destruction of information or information systems. Although the Bureau performed risk assessments

⁴ Federal Information Security Management Act of 2002, Title III, E-Government Act of 2002, Pub. L. No. 107-347 (Dec. 17, 2002).

for some of the IT components used in the 2004 census test, the documentation was not consistent. For example, documentation of information sensitivity risks (high, medium, and low) for confidentiality, integrity, and availability of information were not consistent and did not always follow Bureau policy. In addition, documents showed different numbers of file servers, firewalls, and even different names of devices. Without complete and consistent risk assessment documentation, the Bureau had limited assurance that it properly understood the security risks associated with the test.

- **The Bureau did not always follow its own risk policies.** FISMA requires the implementation of policies and procedures to prevent and/or mitigate security risks. Although Bureau policies allowed for the waiver of security policies, if appropriate, we noted that such policies were not always followed. For example, a waiver for the test of certain password policies was not properly documented and other system documents were not properly updated to reflect the waiver. As a result, the risk assessment for the 2004 census test did not properly identify the related risks and did not identify any compensating controls to reduce the risk to an acceptable level.

As the Bureau plans future tests and the census itself, it will be important for it to strengthen its IT security risk management practices, ensuring they fully adhere to FISMA requirements and its own IT security policies.

Test Reveals Technical, Training, and Other Challenges in Need of Prompt Resolution

The 2004 test suggests that while certain census initiatives have potential, formidable challenges remain. For example, the HHCs show promise in that enumerators were successful in using them to collect data from nonrespondents and remove late mail returns. Still, they were not street ready as they experienced transmission and memory overload problems. Likewise, automated maps were difficult to use, certain questionnaire items confused respondents, and enumerators did not always follow interview protocols. These problems shed light on issues in need of the Bureau's attention as it develops solutions and incorporates refinements for additional testing in the years ahead.

HHCs Were Effective for Conducting Interviews and Removing Late Mail Returns

The Bureau purchased 1,212 HHCs for the test at a total cost of about \$1.5 million. The devices were sent directly to the two test sites packaged in kits that included a battery, AC adaptor, and modem card for transmitting

data via the telephone. The HHCs were also equipped with a Global Positioning System (GPS), a satellite-based navigational system to help enumerators locate street addresses. The Bureau anticipates the HHCs will allow it to eliminate the millions of paper questionnaires and maps that enumerators need when following up with nonrespondents, thereby improving their efficiency and reducing overall costs.

Because the Bureau had never used HHCs in the decennial census, an important goal of the test was to see whether enumerators could use them for interviewing nonrespondents (see fig. 3). Most workers we observed had little trouble using the device to complete the interviews. In fact, most said they were pleased with the HHC's overall functionality, durability, screen clarity, and the ability to toggle between the questionnaire and taking GPS coordinates.

Figure 3: An Enumerator Using an HHC for Nonresponse Follow-up



Source: U.S. Census Bureau.

Another important function of the HHC was removing late mail returns from each enumerator's assignment area(s). Between the Georgia and Queens test sites, over 7,000 late mail returns were removed, reducing the total nonresponse follow-up workload by nearly 6 percent.

The ability to remove late mail returns from the Bureau's nonresponse follow-up workload could help save money in that it could eliminate the need for enumerators to make expensive follow-up visits to households that return their questionnaires after the mail-back deadline. Had the Bureau possessed this capability during the 2000 Census, it could have eliminated the need to visit nearly 773,000 late-responding households and saved an estimated \$22 million (based on our estimate that a 1-percentage point increase in workload could add at least \$34 million in direct salary, benefits, and travel costs to the price tag of nonresponse follow-up⁵). Because of the Bureau's experience in 2000, in our 2002 report on best practices for more cost-effective nonresponse follow-up, we recommended, and the Bureau agreed, that it should develop options that could purge late mail returns from its nonresponse follow-up workload.⁶

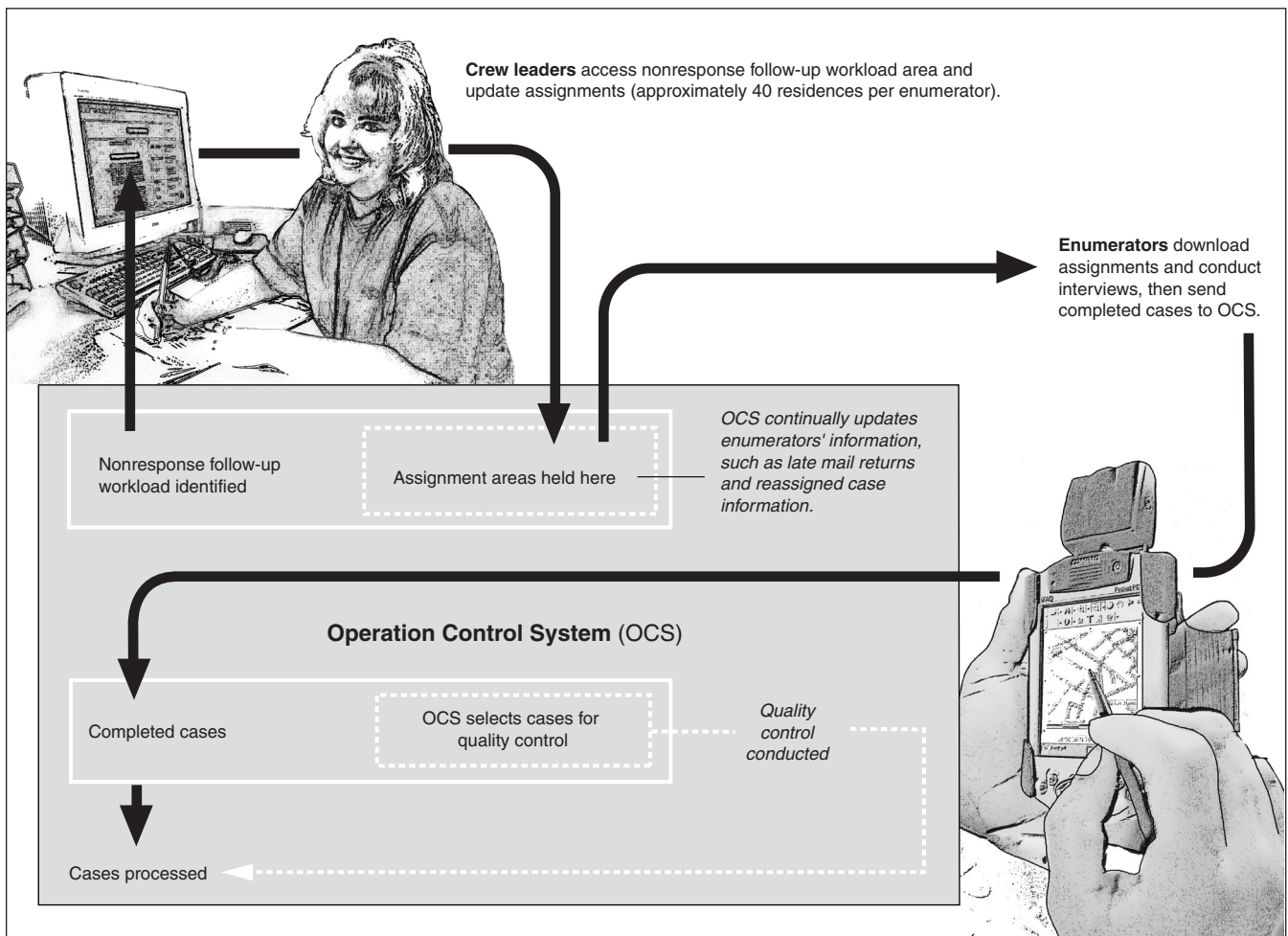
Technical and Training Difficulties Caused HHC Transmission Problems

Each day, enumerators were to transmit completed nonresponse follow-up cases to headquarters and receive assignments, software uploads, or both via a telephone modem (see fig. 4 for a flowchart describing the file transmission process). However, the majority of workers we interviewed had problems doing so, in large part because of technical reasons or because the Bureau's training did not adequately prepare them for the complexity of the transmission procedure, which was a multistep process involving the connection of a battery pack, cables, and other components. As reliable transmissions are crucial to the success of nonresponse follow-up, it will be important for the Bureau to resolve these issues so that the HHCs can be reevaluated in 2006.

⁵ GAO, *2000 Census: Contingency Planning Needed to Address Risks That Pose a Threat to a Successful Census*, [GAO/GGD-00-06](#) (Washington, D.C.: Dec. 14, 1999).

⁶ GAO, *2000 Census: Best Practices and Lessons Learned for More Cost-Effective Nonresponse Follow-up*, [GAO-02-196](#) (Washington, D.C.: Feb. 11, 2002).

Figure 4: Data Transmission Process for Nonresponse Follow-up



Source: GAO analysis of U.S. Census data.

Difficulties began during training when the first transmission was supposed to occur and continued through the remainder of the test. During that first transmission, the Bureau needed to upload a number of software upgrades along with each census worker's first assignment. Many of these transmissions failed because of the volume of data involved. Thus, without cases, the trainees could not complete an important section of on-the-job training. The Bureau acknowledged that these initial problems could have been avoided if the final version of software had been installed on the devices prior to their distribution at training.

Transmission problems persisted throughout nonresponse follow-up. According to the Bureau, during the first 2 weeks of this operation, successful data transmission occurred 80 percent of the time once a connection was made. However, a number of enumerators never even established a connection because of bad phone lines, incorrect passwords, and improper setup of their modems. Other transmission problems were due to the local telecommunication infrastructure at both test sites. For example, in Georgia, older phone lines could not always handle transmissions, while in Queens, apartment intercoms that used phone lines sometimes interrupted connections.

Further, while the transmission rate ultimately increased to 95 percent—roughly the maximum allowed by the technology—that level is still short of the performance level needed for 2010. During the 2000 Census, a 95 percent success rate would have resulted in the failure to transmit around 30,000 completed questionnaires *each day*.

During the test, the Bureau also had to contend with census workers who were “living off the grid”; that is, they only used cellular phones and lacked landlines to transmit and receive data from their homes. While individuals could make alternative arrangements, such as using a neighbor’s telephone, an increasing number of people nationwide in the coming years might give up their landline service to rely on cellular phones, which could be problematic for the Bureau. Bureau officials have noted that all these transmission problems need to be addressed before 2010.

HHCs experienced memory overloads if too many assignment areas were loaded onto them. An assignment area typically contains 40 housing units or cases that are assigned to an enumerator for nonresponse follow-up. The design was to have an entire assignment area transmitted to the HHC even when as few as one case needed follow-up. However, some enumerators’ HHCs became overloaded with too much data, as cases had to be reassigned due to staff turnover, a larger-than-expected number of refusals, and reassignments resulting from language problems. As such, when HHCs became overloaded they would crash and enumerators had to reconfigure them at the local census office, which made them less productive. To the Bureau’s credit, during the test, it was able to work out a solution to avoid overloads by assigning individual cases instead of the entire assignment area to a census worker’s HHC.

Another problem that surfaced during the test was that the HHC’s mapping feature was difficult to use. To contain costs and increase efficiency, the

Bureau expects to replace paper maps with the electronic maps loaded on the HHCs for 2010. However, during the test, enumerators reported that they did not always use the mapping function because it ran slowly and did not provide sufficient information. Instead, they relied on local maps or city directories, and one worker explained that she found it easier to use an Internet mapping service on her home computer to prepare for her route.

Without the Bureau's maps, enumerators might not properly determine whether a housing unit was located in the Bureau's geographic database. This verification is important for ensuring that housing units and the people who reside in them are in the correct census block, as local and state jurisdictions use census population figures for congressional redistricting and allocating federal funds.

Enumerators were also unable to use the HHCs' "go back" function to edit questionnaires beyond a certain point in the interview. In some cases, this led to the collection of incorrect data. For example, we observed one worker complete half an interview, and then discover that the respondent was providing information on a different residence. After the census worker entered the number of residents and their names, the "go back" function was no longer available and as a result that data could not be deleted or edited. Instead, the worker added information in the "notes section" to explain that the interview had taken place at the wrong household. However, Bureau officials told us that they had not planned to review or evaluate these notes and were not aware that such address mix-ups had been documented in the notes section.

To the extent address mix-ups and other inconsistencies occur and are not considered during data processing, accuracy could be compromised. In earlier censuses when the Bureau used paper questionnaires, if workers made mistakes, they could simply erase them or record the information on new forms. As mistakes are inevitable, it will be important for the Bureau to ensure that the HHCs allow enumerators to edit information, while still maintaining the integrity of the data.

Bureau Needs to Review Format of Coverage Improvement and Race/Ethnicity Questions

We found that questions designed to improve coverage and better determine race and ethnicity were awkward for enumerators to ask and confusing for respondents to answer. Consequently, enumerators sometimes did not read the questions exactly as worded, which could adversely affect the reliability of the data collected for these items, as well as the Bureau's ability to evaluate the impact of the revised questions. Our

observations also highlight the importance of ensuring that workers are trained to follow interview protocols; this issue will be discussed later in this report.

Coverage Improvement

While the Bureau attempts to count everyone during a census, inevitably some people are missed and others are counted more than once. To help ensure that the Bureau properly counts people where they live, the Bureau revised and assessed its residency rules for the 2004 census test. For example, under the residence rules, college students should be counted at their campus addresses if they live and stay there most of the time. The Bureau also added two new coverage questions aimed at identifying household residents who might have been missed or counted in error (see fig. 5 for coverage questions).

read the two coverage questions. However, during our field visits we noted that they did not consistently use the flashcards, preferring to summarize them instead. Likewise, enumerators did not always ask the new coverage questions as written, sometimes abbreviating or skipping them altogether. A frequent comment from the workers we spoke with was that the two new coverage questions were awkward because the questions seemed redundant. Indeed, one census worker said that he asked the overcount and undercount questions more times than not, but if people were in a hurry, he did not ask the questions. During one of these hurried interviews, we observed that the census worker did not ask the questions and simply marked “no” for the response.

Race and Ethnicity Questions

Collecting reliable race and ethnicity data is an extremely difficult task. Both characteristics are subjective, which makes accurate measurement problematic. In 2003, the Bureau tested seven different options for formatting the race and ethnic questions, and selected what it thought was the optimal approach to field test in 2004. The Bureau planned to examine respondent reaction to the new race and Hispanic origin questions by comparing responses collected using the paper questionnaire to responses recorded on the HHCs during nonresponse follow-up.

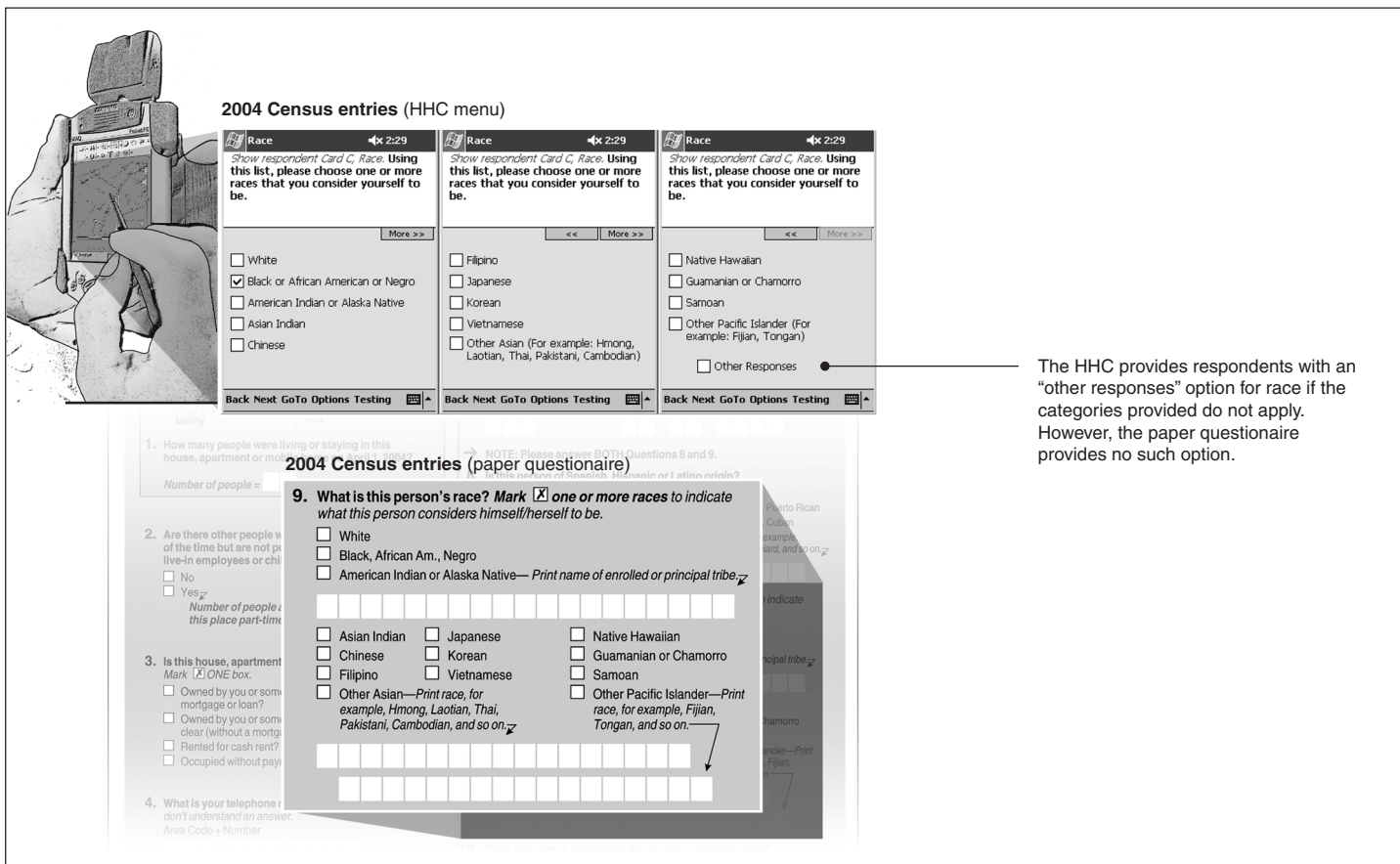
One change the Bureau planned to analyze was the removal of the “some other race” write-in option from the questionnaire. In 2000, the Bureau found that when given this option, respondents would check off “some other race,” but did not always write in what their race was. Thus, in the 2004 test, the Bureau wanted to assess respondents’ reaction to the removal of the “some other race” write-in option. Specifically, the Bureau wanted to see whether respondents would skip the item or select from one of the other options given.

However, we found that the Bureau formatted the race question on the paper questionnaire differently from the question on the HHC. As shown in figure 6, on the paper version, there is not a category for another race other than those categories listed, thus forcing respondents to select a category or skip the question entirely.

This contrasts with the HHCs where, if respondents do not fit into one of the five race categories, the questionnaire format allows them to provide an “other” response and enumerators can record their answers. In fact, the HHC requires enumerators to record a response to the race question and will not allow the interview to continue until a response is entered. As a result, the data recorded by the two questionnaire formats are not

comparable as they could produce different data depending on the data collection mode.

Figure 6: Race and Ethnicity Categories on the HHCs Were Formatted Differently From the Paper Questionnaires



Source: U.S. Census Bureau.

According to the Bureau, it formatted the paper version of the race question differently from the HHC version because it considered the "other" response option on the HHC a respondent comment and not a write-in response. Nevertheless, if the Bureau's purpose is to measure respondent reaction to eliminating the write-in option, it is uncertain what conclusions the Bureau will be able to draw given that this option, even though in the form of a comment, is still available to the respondent during the nonresponse follow-up interview.

As was the case with the coverage measurement question, enumerators at both test locations did not always follow proper interview procedures because they felt the questions were awkward to ask and confused respondents. For example, some workers did not use the flashcards designed to guide respondents in selecting categories for their race and ethnicity and to ensure data consistency. One census worker said that rather than use the flashcards or ask the questions, he might “eyeball” the race and ethnicity. Another worker said that most people laughed at the Spanish, Hispanic, or Latino origin question and she had complaints about the wording of this question. A third census worker noted that he was “loose with the questions” because he could pose them better. Like lapses to the coverage improvement procedures for the 2004 census test, deviating from the interview procedures for the new race and ethnicity questions may affect the reliability of the data and the validity of the Bureau’s conclusions concerning respondent reaction to these questions.

Since the 2004 census test, the 2005 Consolidated Appropriations Act⁷ required that the Bureau include “some other race” as a category when collecting census data on race identification. Consequently, the Bureau said it will include this category on all future census tests and the 2010 Census itself. Thus, while research into eliminating the “some other race” category is now moot, it will still be important for the Bureau to have similar formats for the HHCs and paper questionnaires so that similar data can be captured across modes. Likewise, it will be important for the wording of those questions to be clear and for enumerators to follow proper procedures during interviews.

New Procedures Should Help Reduce Duplicate Enumerations of Group Quarter Residents, but Other Challenges Remain

As noted previously, under its residence rules, the Bureau enumerates people where they live and stay most of the time. To facilitate the count, the Bureau divides residential dwellings into two types: housing units, such as single-family homes and apartments, and group quarters, which include dormitories, prisons, and nursing homes.

The Bureau tested new group quarters procedures in 2004 that were designed to address the difficulties the Bureau had in trying to identify and count this population group during the 2000 Census. For example,

⁷ Consolidated Appropriations Act, 2005, Pub. L. No. 108-447, Div. B, Title II, Dec. 8, 2004.

communities reported instances where prison inmates were counted in the wrong county and residents of college dormitories were counted twice.

One refinement the Bureau made was integrating its housing unit and group quarter address lists in an effort to avoid counting them once as group quarters and again as housing units, a common source of error during the 2000 Census. Census workers were then sent out to verify whether the dwellings were in fact group quarters and, if so, to classify the type of group quarter using a revised “other living quarters facility” questionnaire.

A single address list could, in concept, help reduce the duplicate counting that previously occurred when the lists were separate. Likewise, we observed that census workers had no problems using the revised facility questionnaire and accompanying flashcard that allowed the respondent to select the appropriate type of living facility. This new procedure addresses some of the definitional problems by shifting the responsibility for defining the group quarter type from the Bureau to the respondent, who is in a better position to know about the dwelling.

Another change tested in 2004 was the classification of group homes, which in 2000 was a part of the group quarter inventory. Group homes are sometimes difficult for census workers to spot because they often look the same as conventional housing units (see fig. 7). As a result, they were sometimes counted twice during the 2000 Census—once as a group quarter, and once as a housing unit. For the 2004 test, the Bureau decided to treat group homes as housing units and include them in the housing unit list.

Figure 7: Group Homes Could Resemble Conventional Houses



Source: GAO, 2004.

Early indications from the Bureau suggest that including group homes as housing units, whereby they receive a short-form questionnaire in the mail, may not work. According to the Bureau, the format of the short form is not well suited to group home residents. For example, the questionnaire asks for the “name of one of the people living or staying here who owns or rents this place.” Since the state or an agency typically owns group homes, these instructions do not apply. The Bureau stated that it plans to reassess how it will identify and count people living in group homes.

We identified other problems with the Bureau’s group quarters validation operation during the 2004 census test. For example, we were told that census workers were provided maps of the areas they were assigned but needed maps for adjoining areas so that they could more accurately locate the physical location of the group quarters. In Georgia, where workers used address data from the 2000 Census, the crew leader explained that approximately one-third of all the addresses provided were incorrectly spotted on maps and had to be redone. They also lacked up-to-date

instructions—for example, they did not know that they were to correct addresses rather than just delete them if the addresses were wrong. Further, census workers said that scenarios in the manual and classroom training were based on perfect situations; thus, they did not provide adequate training for atypical settings or when problems arose.

The Bureau Should Rethink Its Approach to Training Enumerators

The success of the census is directly linked to the Bureau's ability to train enumerators to do their jobs effectively. This is a tremendous task given the hundreds of thousands of enumerators the Bureau needs to hire and train in just a few weeks. Further, enumerators are temporary employees, often with little or no prior census experience, and are expected, after just a few days of training, to do their jobs with minimal supervision, under sometimes difficult and dangerous conditions. Moreover, the individuals who train enumerators—crew leaders—are often recent hires themselves, with little, if any, experience as instructors. Overall, few, if any, organizations face the training challenges that confront the Bureau with each decennial population count.

To train the 1,100 enumerators who conducted nonresponse follow-up for the 2004 test, the Bureau employed essentially the same approach it has used since the 1970 Census: crew leaders read material word-for-word from a training manual to a class of 15 to 20 students. The notable exception was that in transitioning from a paper questionnaire to the HHCs, the Bureau lengthened the training time from 3 days to 5 days. However, given the demographic and technological changes that have taken place since 1970, the Bureau might want to explore alternatives to this rigid approach.

As noted earlier, during nonresponse follow-up, enumerators experienced a variety of problems that could be mitigated through improved training. The problems included difficulties setting up equipment to transmit and download data; failure to read the coverage and race/ethnicity questions exactly as worded; and not properly using the flashcards, which were designed to help respondents answer specific questions.

Most of the shortcomings related to training that we observed during the test were not new. In fact, the Bureau had identified these and a number of other training weaknesses in its evaluation of the 2000 Census, but it is clear they have not been fully resolved. Thus, as the Bureau plans for the 2010 Census, it will be important for it to resolve long-standing training problems as well as address new training issues, such as how best to teach

enumerators to use the HHCs and their associated automated processes. Our observations of the test point to specific options the Bureau might want to explore. They include (1) placing greater emphasis on the importance of following prescribed interview procedures and reading questions exactly as worded; (2) supplementing verbatim, uniform training with modules geared toward addressing the particular enumeration challenges that census workers are likely to encounter at specific locales; and (3) training on how to deal with atypical situations or respondent reluctance.

To help evaluate its future training needs, the Bureau hired a contractor to review the training for the 2004 test and recommend actions for improving it. From GAO's work on assessing agencies' training and development efforts, we have developed a framework that can also help in this regard.⁸ Though too detailed to discuss at length in this report, highlights of the framework, and how they could be applied to census training, include:

1. performing proper front-end analysis to help ensure that the Bureau's enumerator training is aligned with the skill and competencies needed to meet its field data collection requirements and work processes and that the Bureau leverages best practices and lessons learned from training enumerators and from past experience;
2. identifying specific training initiatives that in conjunction with other strategies, improve enumerators' performance and help the Bureau meet its goal of collecting high-quality data from nonrespondents;
3. ensuring effective and efficient delivery of training that reinforces new and needed competencies, skills, and behaviors without being wedded to past, and perhaps outmoded, methods; and
4. evaluating the training to ensure it is addressing known skill and competency weaknesses through such measures as assessing participant reactions and changes in enumerators' skill levels and behaviors.

⁸ GAO, *Human Capital: A Guide for Assessing Strategic Training and Development Efforts in the Federal Government*, [GAO-04-546G](#) (Washington, D.C.: March 2004).

Readiness Will Be Critical for Future Tests

Several key features of the 2004 test were not test ready; that is, they were not fully functional or mature when they were employed at the test sites. This is a serious shortcoming because it hampered the Bureau from fully evaluating and refining the various census-taking procedures that will be used in subsequent tests and the actual census in 2010. Further, to the extent these features were integrated with other operations, it impeded the Bureau from fully assessing those associated activities as well.

Our work, and that of the Department of Commerce Inspector General,⁹ identified the following areas where the Bureau needed to be more prepared going into the test:

- The HHCs crashed, in part, because earlier testing did not identify software defects that caused the download of more data to the HHCs than their memory cards could hold.
- Transmission failures occurred during enumerator training, in part, because the HHCs were shipped without the latest version of needed software. Although the Bureau ultimately provided the latest software after several weeks, the upgraded version was unavailable for training field operations supervisors and crew leaders and for the initial enumerator training.
- According to the Department of Commerce Inspector General, the Bureau finalized the requirements for the new group quarter definitions too late for inclusion in group quarters training manuals. Consequently, the training lacked certain key instructions, such as how to categorize group homes.

The Bureau experienced other glitches during the test that with better preliminary testing or on-site dry runs, might have been detected and possibly addressed before the test started. These included the slow start-up of the HHC's mapping function, and the tendency for apartment house intercoms to interrupt transmissions.

An important objective of any type of test is to identify what is working and where improvements are needed. Thus, it should not be surprising, and, in

⁹U.S. Department of Commerce, Office of Inspector General, *Improving Our Measure of America: What the 2004 Census Test Can Teach Us in Planning for the 2010 Decennial Census*, OIG-16949 (Washington, D.C.: September 2004).

fact, should be expected and commended, that shortcomings were found with some of the various activities and systems assessed during the 2004 test. We believe that the deficiency is not the existence of problems; rather it is the fact that several components were incomplete or still under development going into the test, which made it difficult for the Bureau to gauge their full potential. The Bureau had a similar experience in the dress rehearsal for the 2000 Census, when, because a number of new features were not test ready, the Bureau said it could not fully test them with any degree of assurance as to how they would affect the head count.

Because of the tight time frames and deadlines of the census, the Bureau needs to make the most of its limited testing opportunities. Thus, as the Bureau plans for the next field test in 2006 and the 2008 dress rehearsal, it will be important for the Bureau to ensure the various census operations are fully functional at the time of the test so they can be properly evaluated.

Conclusions

The Bureau is well aware that a successful enumeration hinges on early research, development, testing, and evaluation of all aspects of the census design. This is particularly true for the 2010 Census for which, under its current plan, the Bureau will be relying on HHCs and other methods and technologies that (1) have never been used in earlier censuses and (2) are mission critical. Consequently, the 2004 test was an important milestone in the 2010 life cycle because it demonstrated the fundamental feasibility of the Bureau's basic design and allows the Bureau to advance to the next and more mature phase of planning and development.

Nevertheless, while the test revealed no fatal flaws in the Bureau's approach, the results highlighted serious technical, training, methodological, and procedural difficulties that the Bureau will need to resolve. Since one of the purposes of testing is to determine the operational feasibility of the census design, it is not surprising that problems surfaced. However, looking toward the future, it will be critical for the Bureau to diagnose the source of these challenges, devise cost-effective solutions, and integrate refinements and fixes in time to be assessed during the next field test scheduled for 2006. It will also be important for Congress to monitor the Bureau's progress as it works to resolve these issues.

Recommendations for Executive Action

To facilitate effective census planning and development, and to help the Bureau achieve its key goals for the census—reduce risks, improve accuracy, and contain costs, we recommend that the Secretary of Commerce direct the Bureau to take the following eight actions:

- Analyze the impact that HHCs and the targeted second mailing had on cost savings and other Bureau objectives.
- Ensure the Bureau's IT security practices are in full compliance with applicable requirements, such as the FISMA, as well as its own internal policies.
- Enhance the reliability and functionality of HHCs by, among other actions, (1) improving the dependability of transmissions, (2) exploring the ability to speed up the mapping feature, (3) eliminating the causes of crashes, and (4) making it easier for enumerators to edit questionnaires.
- Define specific, measurable performance requirements for the HHCs and other census-taking activities that address such important measures as productivity, cost savings, reliability, durability, and test their ability to meet those requirements in 2006.
- Review and test the wording and formatting of the coverage and race/ethnicity questions to make them less confusing to respondents and thus help ensure the collection of better quality data, and ensure they are formatted the same way on both the HHC and paper versions of the census form.
- Develop a more strategic approach to training by ensuring the curriculum and instructional techniques (1) are aligned with the skills and competencies needed to meet the Bureau's data collection requirements and methodology and (2) address challenges identified in the 2004 test and previous censuses.
- Revisit group quarter procedures to ensure they allow the Bureau to best locate and count this population group.
- Ensure that all systems and other census-taking functions are as mature as possible and test ready prior to their deployment for the 2006 test, in part by conducting small-scale, interim tests under the various

conditions and environments the Bureau is likely to encounter during the test and actual enumeration.

Further, to ensure the transparency of the census-planning process and facilitate Congressional monitoring, we also recommend that the Secretary of Commerce direct the Bureau to regularly update Congress on the progress it is making in addressing these and any other challenges, as well as the extent to which the Bureau is on track for meeting the overall goals of the 2010 Census.

Agency Comments and Our Evaluation

The Under Secretary for Economic Affairs at the Department of Commerce forwarded us written comments from the Census Bureau on a draft of this report on December 20, 2004, which are reprinted in appendix I. The Bureau noted that the 2004 test was its first opportunity to assess a number of the new methods and technologies under development for 2010, and emphasized the importance of a sustained, multiyear planning, testing, and development program to its census modernization effort.

The Bureau generally agreed with seven of our nine recommendations, and described the steps it was taking to address our concerns. The Bureau also provided additional context and clarifying language and we have added this information to the report where appropriate.

Specifically, the Bureau generally agreed with our recommendations relating to improving IT security practices, the reliability of the HHCs, training, testing, and enumeration procedures—and reported it was already taking a number of steps to address our concerns. We commend the Bureau for recognizing the risks and challenges that lie ahead and taking action to address them. We will continue to monitor the Bureau's progress in resolving these issues and update Congress on a regular basis.

At the same time, the Bureau took exception to our recommendations to (1) analyze the impact that HHCs and targeted second mailings had on cost savings and other Bureau objectives, and (2) define specific, measurable performance requirements for the HHCs and other census-taking activities and test their ability to meet those requirements in 2006. With respect to the first recommendation, the Bureau noted that it did not establish cost-savings and other impacts as test objectives, in part, because the Bureau believes that the national sample mail test that it conducted in 2003 provided a better method for determining the boost in response rates that could accrue from a second mailing. The Bureau maintains that analyzing

the impact of the second mailing would provide it with no more information beyond what it has already established from the 2003 test and would be of little value.

We believe this recommendation still applies because it will be important for the Bureau to assess the impact of the targeted second mailing on other Bureau objectives. As we noted in the report, the Bureau included the HHCs and targeted second mailing in the 2010 Census design, in part, to reduce staff, improve productivity, and control costs. Further, as we also note in the report, the feasibility of a targeted second mailing is an open question. Thus, information on the degree to which the HHCs and second mailing contribute to these key goals could help inform future budget estimates, investment and design decisions, as well as help refine future census tests. In short, the purpose of the analysis we recommend would not be to see whether these features of the 2010 Census will produce cost-savings, but the extent of those savings and the impact on other Bureau objectives.

With respect to the second recommendation, the Bureau noted that it had “baseline assumptions” about productivity, cost-savings, and other measures for the 2004 Census test and that a key objective of the test was to gather information to help refine these assumptions. According to the Bureau, this will also be a key objective of the 2006 Census Test, although its performance goal will not be whether it meets specific measures. Instead, the Bureau intends to focus on successfully collecting information to further refine those assumptions. As a result, the Bureau believes the 2006 test will not be a failure if HHC productivity is not achieved, but that it will be a failure if productivity data are not collected.

The Bureau’s position is inconsistent with our recommendation which we believe still applies. As noted in the report, we call on the Bureau to define measurable performance requirements for the HHCs as well as take the next step and assess whether the HHCs can meet those requirements as part of the 2006 test. This information is essential because it will help the Bureau gauge whether HHCs can meet its field data collection needs in 2010. Should the HHCs fail to meet these pre-specified performance requirements during the 2006 test, the Bureau would need to rethink how it employs these devices in 2010.

As agreed with your offices, unless you release its contents earlier, we plan no further distribution of this report until 30 days from its date. At that

time, we will send copies of this report to the Secretary of Commerce and the Director of the U.S. Census Bureau. Copies will be made available to others on request. This report will also be available at no charge on GAO's home page at <http://gao.gov>. Please contact me at (202) 512-6806 or daltonp@gao.gov or Robert Goldenkoff, Assistant Director, at (202) 512-2757 or goldenkoffr@gao.gov if you have any questions. Key contributors to this report were Tom Beall, David Bobruff, Betty Clark, Robert Dacey, Richard Donaldson, Elena Lipson, Ronald La Due Lake, Robert Parker, Lisa Pearson, and William Wadsworth.



Patricia A. Dalton
Director
Strategic Issues

Comments from the Department of Commerce



UNITED STATES DEPARTMENT OF COMMERCE
The Under Secretary for Economic Affairs
Washington, D.C. 20230

DEC 20 2004

Ms. Patricia A. Dalton
Director, Strategic Issues
U.S. Government Accountability Office
Washington, DC 20548

Dear Ms. Dalton:

The U.S. Department of Commerce appreciates the opportunity to comment on the U.S. Government Accountability Office draft report entitled *2010 Census: Basic Design Has Potential, but Remaining Challenges Need Prompt Resolution*. The Department's comments on this report are enclosed.

The Department of Commerce and the U.S. Census Bureau stand ready to update the Congress at any time on these matters.

Sincerely,


Kathleen B. Cooper

Enclosure



Appendix I
Comments from the Department of
Commerce

Comments from the U.S. Department of Commerce, U.S. Census Bureau,
Regarding the U.S. Government Accountability Office (GAO) Draft Report Entitled
2010 Census: Basic Design Has Potential, but Remaining Challenges Need Prompt
Resolution

Thank you for the opportunity to comment on your draft report, *2010 Census: Basic Design Has Potential, but Remaining Challenges Need Prompt Resolution* (GAO-05-9), which focuses on our 2004 Census Test. We would like the final report to include the following statement about the importance and role of the 2004 Census Test in our overall efforts. In addition, we have provided comments on each of your nine recommendations for executive action.

Importance and Role of the 2004 Census Test

The 2004 Census Test was the Census Bureau's first opportunity to field test a number of the new methods and technologies being developed for the 2010 census. In response to the lessons of Census 2000, and in striving to better meet this Nation's ever-expanding needs for social, demographic, and geographic information, the U.S. Department of Commerce and the U.S. Census Bureau have developed a multiyear plan to completely modernize and reengineer the 2010 Census of Population and Housing. This reengineering effort has four major goals: improve the relevance and timeliness of census long-form data; reduce operational risk; improve the accuracy of census coverage; and contain costs.

A sustained, multiyear, integrated program for planning, testing, and developing such a short-form-only census for 2010 is a key component of our reengineering effort. The Census Bureau appreciates the support we have received for these efforts from the Administration, the Congress, and other key stakeholders.

Specific Comments About the Report's Recommendations for Executive Action

GAO Recommendation 1: Analyze the impact that HHCs and targeted second mailing had on cost savings and other Bureau objectives.

Response:

Regarding the impact of the targeted second mailing on cost savings, the Census Bureau did not establish this as a test objective for the 2004 Census Test for a number of reasons. The Census Bureau believes that a national sample mail test provides a better vehicle for assessing the gain in response rates from a second mailing that we could expect to realize during the 2010 census. Thus, as part of the 2003 National Census Test, we included objectives and conducted thorough analyses of the gains in response associated with targeted replacement mailings. Those results confirmed earlier findings that, nationally, we could expect a 7-to-10 percentage point gain in overall response, based on the delivery of a targeted replacement mailing.

Appendix I
Comments from the Department of
Commerce

Subsequent to the 2003 National Census Test, we have had, and continue to have, extensive consultations with the printing industry to identify viable technical solutions to our replacement mailing objective for the 2010 census. The key to our ability to consult directly with industry on potential solutions for a successful 2010 replacement mailing strategy was an agreement between the Census Bureau and the Government Printing Office, which allowed us to conduct such consultations with industry without prejudice to future contracting opportunities. These discussions with the printing industry have provided us with a reasonable degree of assurance that a targeted replacement mailing strategy is, indeed, a viable option for the 2010 census. As a result of this consultation, we plan to evaluate in the 2005 National Census Test, from the standpoint of public reaction as measured by mail response, at least two different replacement mailing package designs that could be printed and delivered within the very tight time constraints that we will face in the 2010 census. Such production solutions had not been identified in time to evaluate them in the 2004 Census Test. Thus, establishing an objective regarding replacement mailing for that test would have provided no more intelligence beyond what we had already established from the 2003 National Census Test—that is, that such strategies increase mail response. Indeed, our operational assessment of the 2004 Census Test once again demonstrated that the targeted replacement mailing strategy increased mail response by about 7 percentage points in the Queens test site. However, since the solution implemented in that test (namely, using an in-house address imaging solution) is not one that is viable given the large volumes and compressed time schedule that we will face during the 2010 census, devoting resources to a formal evaluation of it now would be of little value.

Looking toward the 2005 National Census Test, the 2006 Census Test, and the 2008 Dress Rehearsal, we are continuing our research and exploration of an array of issues related to a successful implementation of a targeted replacement mailing strategy. We will look at the timing and process for identifying the replacement universe, file exchange and security, data capture, and other such issues that need to be addressed within a total systems approach for achieving our objectives. We believe that these research efforts, when combined with the future testing and demonstration vehicles, will provide us an excellent opportunity to realize the substantial cost savings associated with a targeted replacement mailing strategy in the 2010 census.

Another benefit of a successful mailing strategy, combined with the use of hand-held computers (HHCs), is the ability to remove late mail returns from the nonresponse follow-up work load. During Census 2000, we received over 3 million mail returns after we had identified and assigned follow-up visits for the nonresponding households. For the 2010 census, we envision that the use of HHCs will enable us to efficiently remove, on a daily basis, a substantial proportion of these late mail returns from enumerator assignments. Again, our operational assessment of the 2004 Census Test suggests that we removed a substantial percent of the late mail returns from the enumerator assignments before follow-up visits occurred. Given the potential impact that this process has on total follow-up costs, we continue to conduct simulations, using actual Census 2000 data, to factor this into our cost parameters for the 2010 census.

GAO Recommendation 2: Ensure the Bureau's IT security practices are in full compliance with applicable requirements, such as the FISMA, as well as its own internal policies.

Response:

The GAO noted the following areas of concern in their report, and the Census Bureau is developing procedures to address these issues in a manner consistent with program objectives and overall risk to our information and information systems.

IT inventory was not complete: The IT Security Office is working with the Computer Services Division and the Telecommunications Office, as well as the program areas, to ensure that all IT systems identified in the Census Bureau sensitive system inventory have complete and accurate inventories. We are accomplishing this by using information collected by the Bowie Computer Center (BCC) during weekly IT discovery operations. The IT Security Office receives these reports concerning its security documentation. IT security and system owners reconcile inventories. Discrepancies are investigated with the program areas to ensure that the correct information is provided in the documentation. System owners are ultimately responsible for maintaining an accurate inventory of their systems, and the IT Security Office is working with Division Security Officers (DSO) to ensure they are aware of this responsibility and are taking appropriate actions.

There was not sufficient evidence that the Bureau assessed all of the devices used in the test for vulnerabilities, or that it corrected previously identified problems: The lack of evidence that the devices were assessed for vulnerabilities was due to high level, rather than detailed and accurate, security documentation. Security documentation, if completed correctly, will provide the appropriate level of evidence that devices were tested for vulnerabilities and what corrective actions, if any, were taken to correct or mitigate the resulting risks. This was not clearly documented, and the GAO reviewer was unable to verify what, if anything, was done in this regard. The IT Security Office is working with program areas (through the DSO program and meetings with system owners or their representatives) to ensure they understand the System Development Life Cycle, including the testing requirements and how to document them correctly.

Assessments were not consistent: Since the GAO review was completed, the IT Security Office has required that an updated risk assessment for each system be prepared and submitted as part of the IT security documentation package. The IT Security Office is currently reviewing these documents for consistency and compliance with NIST 800-30, "Risk Management Guide for Information Technology Systems." In addition, the IT Security Office is looking at the assessments to ensure that, where appropriate, assessment information from other systems is addressed. This was addressed specifically by the GAO review, in that

**Appendix I
Comments from the Department of
Commerce**

information from the IT infrastructure assessments was not communicated to the program areas for consideration and inclusion in their risk-management documentation.

The Bureau did not always follow its own risk policies: The IT Security Office is scheduling formal reviews to ensure that documentation is updated in a timely manner when changes are made to a system. This would include waiver information.

The IT Security Office has purchased an Enterprise tool to assist in the documentation, certification, and accreditation process. The tool, "Xacta," when fully implemented, will ensure that all documentation is consistent across the Census Bureau and that, where applicable, information that pertains to more than one system is made available to the other appropriate area for inclusion in its documentation, as well.

Ensure the Bureau's IT security practices are in full compliance with applicable requirements, such as the FISMA, as well as its own internal policies: The Census Bureau agrees with the draft GAO report that it needs to improve the Census Bureau's IT security practices to ensure the organization is in full compliance with applicable Federal Information Security Management Act (FISMA) and internal policies. The Census Bureau has already sought to improve the IT Security posture of the Census Bureau for the 2006 Census Test in several tangible ways. First, the Technologies Management Office (TMO) hired two IT security contractors from the Xacta Corporation to overhaul the organization's IT security program. We also are implementing the industry standard Xacta IA Manager tool. This software suite enables the organization to meet all government standards for IT security requirements, manage the organization's risk-assessment approaches in a more thorough fashion, and ensure consistency.

The acquisition of the security contracting experts and the use of a comprehensive security and risk-assessment tool help us ensure that the organization will address the FISMA and the NIST SP 800-30, SP 800-26 guidelines, and Census Bureau internal IT Security guidelines more thoroughly. In addition to the tools and contractors, we are engaging in more thorough risk assessments and contingency planning, and are evaluating and implementing the findings of the Inspector General and GAO reviews conducted on the 2004 Census Test. These reviews have helped focus the organization's attention on critical areas of concern.

GAO Recommendation 3: Enhance the reliability and functionality of HHCs by, among other actions, (1) improving the dependability of transmissions, (2) exploring the ability to speed up the mapping feature, (3) eliminating the causes of crashes, and (4) making it easier for enumerators to edit questionnaires.

Appendix I
Comments from the Department of
Commerce

Response:

In general, we agree with the draft GAO recommendations to enhance the reliability and functionality of the HHCs. The TMO carefully examined the results of the reliability and functionality of HHCs during the 2004 Test Census. After evaluating the transmission process, we are implementing several improvements. The first improvement, which was actually implemented late in the 2004 Census Test, was to upgrade the version of the Afaria software system. This significantly reduced transmission errors later in the 2004 Census Test. We expect that this improved performance will continue during the 2006 Census Test, although it is uncertain if dial-up technology can exceed a 95 percent success rate.

A second improvement was the simplification of the transmission process. During the 2004 Census Test, transmissions were part of the Assignment Management System (AMS), which required the enumerator to go into AMS and then transmit their data. The 2006 Census Test software has been arranged in such a way that transmissions are accomplished outside of AMS. In addition, more programming was added to decode arcane transmission error messages. This enables the enumerators and technical support staffs to better understand HHC communication and transmission problems, so that remedies can be applied more accurately and quickly.

To address speeding up the HHC mapping features, we obtained compression software that provides better performance for map displays for the 2006 Census Test. In addition, the HHC procured for the 2006 Census Test has a faster processor than the one used in the 2004 Census Test. The Census Bureau also procured a new generation of SD memory cards, which we anticipate will improve performance of applications between the actual HHC device and the memory storage cards.

To eliminate the causes of crashes, we have taken several approaches. First, we upgraded the operating system that was used in the 2004 Census Test. The previous version of the operating system was less robust and had documented system instabilities that have been addressed by the newer operating system. Second, we streamlined our in-house software design to fix issues found in the 2004 Census Test. We removed a third-party product, which allows the developers greater programming flexibility and redesigned the database on the HHC to make it more efficient. Lastly, we have initiated a comprehensive integration work group that consists of all Census Bureau development and deployment entities. This work group already has begun to work out processes to more effectively share device memory and work out software bugs in a collaborative fashion.

The Census Bureau is unclear about the meaning of the GAO's use of the terminology "editing questionnaires and how we can make it easier for the enumerator." We will address this upon receiving clarification of this statement.

Finally, the Census Bureau has sought to improve the testing process by ensuring that sufficient time is planned into our schedules for more integrated testing (in addition to

Appendix I
Comments from the Department of
Commerce

our standard unit testing and user acceptance testing). We have improved our test plans, test cases, and test reporting by including new performance metrics. Finally, the TMO testing area has added a senior IT specialist to the team, whose sole focus is the 2006 Census Test.

GAO Recommendation 4: Define specific, measurable performance requirements for the HHCs and other census-taking activities that address such important measures as productivity, cost savings, reliability, durability, and test their ability to meet those requirements in 2006.

Response:

We had baseline assumptions about these measures for the 2004 Census Test, and a key objective of the test was to gather information to help refine these assumptions for the 2006 Census Test and for 2010. This also will be a key objective of the 2006 Census Test, but the performance goal will not be whether we meet specific measures based on the revised assumptions; rather, we will focus on successfully collecting information to further refine those assumptions. For example, the 2006 Census Test will not be a failure if we do not achieve the assumed productivity rate for HHCs, but it will be a failure if we do not capture such HHC productivity data in order to refine our assumption for 2010.

GAO Recommendation 5: Review and test the wording and formatting of the coverage and race/ethnicity questions to make them less confusing to respondents and thus help ensure the collection of better quality data, and ensure they are formatted the same way on both the HHC and paper versions of the census form.

Response:

The Census Bureau concurs with the spirit, if not the details, of this recommendation. Following the 2004 Census Test, we conducted extensive cognitive tests of both the coverage probes and residence rules presentation, as well as the race and ethnic-origin questions. Further, the input to these cognitive tests was refined based on close consultation with our advisory committees, the National Academy of Sciences, and, in the case of the race and ethnicity questions, by consultation with demographic and social science experts.

It is important to underscore that, as mandated by Congress, we intend to include a "Some Other Race" (SOR) category in future decennial census collection efforts. The SOR response option is the mechanism through which we obtain other race information. In addition, we are carefully analyzing the results of our cognitive tests to inform decisions about other question wording and design alternatives to be tested in the 2005 National Census Test, the major testing vehicle for deciding on the design, content, and construction of these questions for both the 2010 census as well as the American Community Survey.

Appendix I
Comments from the Department of
Commerce

The Census Bureau disagrees that questions need to be formatted identically between different data collection modes. Rather, we believe, and survey design experts concur, that the ideal goal is to optimize questions for a particular mode to yield similar outcomes across modes. What this means in a practical sense is that question format and presentation in an electronic mode should not necessarily mimic a paper format and presentation. For the 2004 Census Test, one of our key objectives was to ascertain public reaction, as well as the enumerator/respondent interaction, to removing the SOR response category from the census test. Thus, on the mail-out paper questionnaire, the race question did not display the SOR category. And, it is important to note, that respondents could, and indeed did, write in a race response that did not include one of the five displayed race categories.

Similarly, in our design of this question for HHC, we were very much aware that some respondents would name a race response that would not be one of the five categories. To accommodate this, and more importantly, to assess public reaction to the lack of a SOR response category, we quite deliberately designed a notes space on the HHC so that enumerators could record these responses as well as other pertinent information about the enumerator/respondent interaction. Rather than mimic the paper questionnaire design, we designed the HHC in such a way that we could get as much intelligence and information as possible about issues that both respondents and enumerators were experiencing when faced with the removal of the SOR category. The 2004 Census Test was not at all about determining if we obtained similar race distributions between mail respondents and nonrespondents—we already know that these are quite different universes. Rather, it was designed to understand issues associated with the removal of the SOR response during door-to-door enumeration and to make appropriate design decisions based on that intelligence. Now that the Congress has mandated the use of the "SOR" category, some of these issues are moot.

GAO Recommendation 6: Develop a more strategic approach to training by ensuring the curriculum and instructional techniques (1) are aligned with the skills and competencies needed to meet the Bureau's data collection requirements and methodology and (2) address challenges identified in the 2004 test and previous censuses.

Response:

The Census Bureau is exploring ways to improve our training strategy based on the lessons learned from Census 2000 and the 2004 test. However, it is important that the solutions fall within our budget constraints and allow us to deliver consistent training to hundreds of thousands of enumerators across the country. As part of the market research phase of the Field Data Collection Automation (FDCA) contract, we are exploring ways to simplify the hardware, software, and implementation of security, transmission and other requirements that would result in minimizing and simplifying the training needs for automation/technology issues. We also are exploring ways to incorporate technology to assist with training and to identify other ways in which industry can assist with training,

Appendix I
Comments from the Department of
Commerce

such as the “train-the-trainer” concepts. We will use this information to inform the definition of requirements for the FDCA contract.

The Census Bureau recognizes that the training deficiency items listed in the report require improvement. As we prepare for the 2006 test, we are enhancing the training to reinforce the procedural requirements for asking questions exactly as worded and emphasizing the mandatory use of flashcards. We also will incorporate additional training to prepare the enumerators to handle realistic situations encountered in their work.

GAO Recommendation 7: Revisit group quarter procedures to ensure they allow the Bureau to best locate and count this population group.

Response:

The Census Bureau plans to continue to improve procedures and operations to locate and count the population in group quarters. Building upon the lessons learned in the 2004 Census Test, the Census Bureau will:

- (1) Conduct Address Canvassing prior to the Group Quarters Validation Operation (GQV) to correctly locate Other Living Quarters and to minimize duplication between the housing units and group quarters.
- (2) Implement the GQV Operation to determine the classification of the Other Living Quarters and, if it is determined to be a Group Quarters, classify the type of group quarters.
- (2) Refine GQV procedures to clearly instruct listers regarding correcting and deleting addresses.
- (3) Complete the revisions to the definitions for the types of group quarters prior to writing procedures.
- 4) Classify group homes as group quarters, not housing units.

GAO Recommendation 8: Ensure that all systems and other census-taking functions are as mature as possible and test ready prior to their deployment for the 2006 test, in part by conducting small-scale, interim tests under the various conditions and environments the Bureau is likely to encounter during the test and actual enumeration.

Response:

The 2004 Census Test was the Census Bureau’s first opportunity to field test a number of the new methods and technologies being developed for the 2010 census. In response to the lessons of Census 2000, and in striving to better meet this Nation’s ever-expanding

Appendix I
Comments from the Department of
Commerce

needs for social, demographic, and geographic information, the U.S. Department of Commerce and the U.S. Census Bureau have developed a multiyear effort to completely modernize and reengineer the 2010 Census of Population and Housing. This reengineering effort has four major goals: improve the relevance and timeliness of census long-form data; reduce operational risk; improve the accuracy of census coverage; and contain costs.

A sustained, multiyear, integrated program for planning, testing, and developing such a short-form-only census for 2010 is a key component of our reengineering effort. The Census Bureau appreciates the support we have received for these plans from the Administration, the Congress, and other key stakeholders.

The data collection effort for 2010 will take advantage of and build on the American Community Survey and MA/TIGER improvements to contain costs and improve accuracy, while keeping operational risk to a minimum. To make these changes successfully, procedures must be fully tested under census-like conditions and refined well in advance of Census Day.

GAO Recommendation 9: Further, to ensure the transparency of the census-planning process and facilitate Congressional monitoring, we also recommend that the Secretary of Commerce direct the Bureau to regularly update Congress on the progress it is making in addressing these and any other challenges, as well as the extent to which the Bureau is on track for meeting the overall goals of the 2010 Census.

Response:

We concur with this recommendation.

GAO's Mission

The Government Accountability Office, the audit, evaluation and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's Web site (www.gao.gov). Each weekday, GAO posts newly released reports, testimony, and correspondence on its Web site. To have GAO e-mail you a list of newly posted products every afternoon, go to www.gao.gov and select "Subscribe to Updates."

Order by Mail or Phone

The first copy of each printed report is free. Additional copies are \$2 each. A check or money order should be made out to the Superintendent of Documents. GAO also accepts VISA and Mastercard. Orders for 100 or more copies mailed to a single address are discounted 25 percent. Orders should be sent to:

U.S. Government Accountability Office
441 G Street NW, Room LM
Washington, D.C. 20548

To order by Phone: Voice: (202) 512-6000
TDD: (202) 512-2537
Fax: (202) 512-6061

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Web site: www.gao.gov/fraudnet/fraudnet.htm

E-mail: fraudnet@gao.gov

Automated answering system: (800) 424-5454 or (202) 512-7470

Congressional Relations

Gloria Jarmon, Managing Director, JarmonG@gao.gov (202) 512-4400
U.S. Government Accountability Office, 441 G Street NW, Room 7125
Washington, D.C. 20548

Public Affairs

Susan Becker, Acting Manager, BeckerS@gao.gov (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, D.C. 20548

**United States
Government Accountability Office
Washington, D.C. 20548-0001**

**Presorted Standard
Postage & Fees Paid
GAO
Permit No. GI00**

**Official Business
Penalty for Private Use \$300**

Address Service Requested

