

RESEARCH REPORT SERIES  
*(Survey Methodology #2008-4)*

**Report of Cognitive Testing of Privacy-  
and Confidentiality-Related Statements  
in Respondent Materials for the 2010 Decennial:  
Results from Cognitive Interview Pretesting  
with Volunteer Respondents**

Ashley Landreth  
Eleanor Gerber  
Theresa DeMaio

Statistical Research Division  
U.S. Census Bureau  
Washington, DC 20233

Report Issued: March 28, 2008

*Disclaimer:* This paper is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

# FINAL REPORT

---

## Understanding Confidentiality- and Privacy-related Statements in Respondent Materials for the 2010 Decennial: Results from Cognitive Interview Pretesting with Volunteer Respondents



Prepared by: **Ashley Landreth**  
**Eleanor Gerber**  
**Theresa DeMaio**

*U. S. Census Bureau*  
*Statistical Research Division*  
*Center for Survey Methods Research*

Submitted: **September 4, 2007**

# TABLE OF CONTENTS

	Page
<b>I. Executive Summary</b> .....	<b>i</b>
A. Introduction (i)	
<i>Figure 1. Versions of the Confidentiality Paragraph on the Letter's Front</i> (ii)	
<i>Figure 2. Versions of the Confidentiality Paragraph on the Letter's Back</i> (ii)	
B. Findings (iii)	
C. Recommendations (iv)	
<b>II. Introduction</b> .....	<b>1</b>
<b>III. Background</b> .....	<b>3</b>
A. Scope (3)	
B. New Approaches to Conveying the Messages (4)	
C. Past Research & Conceptual Difficulties (6)	
<i>General structure of the messages</i> (7)	
<i>"Statistical purposes"</i> (8)	
<i>Administrative records use</i> (9)	
D. Limitations of the Project, the Messages, and Letter Real Estate (10)	
<b>IV. Method</b> .....	<b>12</b>
A. Research Questions (12)	
B. Pretesting Research Design (12)	
<i>Cognitive interviews</i> (12)	
<i>Protocol</i> (13)	
<i>Materials</i> (13)	
<i>Recruiting</i> (14)	
<i>Respondent characteristics</i> (14)	
<b>V. Findings</b> .....	<b>17</b>
A. Respondents' Interpretations of the Main Messages (17)	
<i>The main confidentiality message</i> (17)	
<i>The "statistical purposes" concept</i> (19)	
<i>Administrative records use message</i> (21)	
<i>Federal law: The "Title 13" concept</i> (30)	
<i>The "72 years" concept</i> (31)	

B. Motivational Messages (33)	
<i>“Protecting your privacy is our highest priority”</i> (33)	
<i>“Sworn for life”</i> (33)	
<i>Penalties</i> (34)	
<i>“Prevents anyone from using them against you”</i> (35)	
<i>Privacy website</i> (35)	
C. Vague Messages (36)	
<i>“Rights and benefits”</i> (36)	
<i>“Oath”</i> (37)	
<i>“Dress rehearsal”</i> (37)	
D. Letter and Language Preferences (38)	
E. Locating Messages on the Back of Letters (39)	
<b>VI. Recommendations</b> .....	<b>40</b>
<b>VII. Conclusion</b> .....	<b>45</b>
A. Answering the Research Questions (45)	
B. Results in Context—A Range of Application (47)	
C. Future Research (48)	
<b>Epilogue</b> .....	<b>49</b>
Figure 1. Versions of the Confidentiality Paragraph on the Letter’s Front .....	<b>5</b>
Figure 2. Versions of the Confidentiality Paragraph on the Letter’s Back .....	<b>5</b>
Figure 3. Simplified Statement on the Census Form .....	<b>6</b>
Figure 4. Reminder Postcard Statement .....	<b>6</b>
Table 1. Randomized Interview Design .....	<b>14</b>
Table 2. Educational Attainment and Geographic Location of Respondents .....	<b>15</b>
<b>Attachment I</b> — Recommended Decennial Cover Letter	
<b>Attachment II</b> — March 15, 2007 Version of the Decennial Cover Letter	

## I. EXECUTIVE SUMMARY

### A. Introduction

At the request of the Decennial Management Division (DMD), staff from the Statistical Research Division (SRD) led an interdivisional team to develop and cognitively pretest new confidentiality language embedded in the mailing package for the 2008 Census Dress Rehearsal. Once developed, the intent was to include these messages in the 2010 Census mail materials. In particular, the team (or “working group”) was given the task of developing new confidentiality- and privacy-related messages for the decennial mailing package, namely, the main confidentiality statement and supporting concepts. The supporting concepts either modify or describe how the data will be treated and/or used (i.e., to produce statistics and/or to study and improve process-oriented production efficiency), and how the data may be augmented by records housed within other federal agencies—concepts commonly referred to as “statistical purposes” and “administrative records use” within the Census Bureau. The Working Group was responsible for identifying the main components of these concepts, in order to communicate them to respondents in plain language, and developing these ideas into a cohesive confidentiality message that would be embedded in a respondent cover letter to the decennial questionnaire.

One aim of the cognitive pretesting research was to determine the best wording for particular “notification” messages the Census Bureau is legally or ethically required (either by way of federal law or Census Bureau policy) to convey to respondents.<sup>1</sup> Another aim of the research was to determine the best location within the letter for particular messages—mainly, we wished to determine whether it was feasible to place some messages on the back of the letter without generating negative reactions from respondents. Toward these goals, the Working Group assembled in the summer of 2006 to draft various versions of these concepts in light of past experimental and qualitative research.

The first two paragraphs of the respondent letter described the 2008 Dress Rehearsal and requested that the form be mailed back promptly. The confidentiality-related content was located in the third paragraph of the respondent letters, and in some cases, continued on the back of the letter where more detailed concepts were expressed. Figures 1 and 2, on the next page, contain excerpts of the three resulting survey cover letters illustrate various approaches used to convey the “statistical purposes” and the “administrative data uses” concepts (See Attachment I at the end of this report for the paragraphs preceding the confidentiality text).

Staff in SRD conducted 50 cognitive interviews between August and November of 2006 with respondents from the Washington, DC and St. Louis, Missouri metro areas, as well as Hawaii. Each respondent reviewed two survey cover letters, a follow-up postcard, and the census form itself. The survey cover letters were the focus of the cognitive research because they alone contained the full suite of messages under investigation, but the follow-up postcard and census form also contained confidentiality-related messages.

---

<sup>1</sup> In “mandatory” data collection efforts, respondents are made aware of certain aspects of the collection and this is typically referred to as “notification” since by definition they do not choose whether to participate. In “voluntary” surveys this same process is referred to as “informed consent.”

**Figure 1. Versions of the Confidentiality Paragraph on the Letter’s Front<sup>2</sup>**

Letter A:	Your answers are confidential. <u>That means the Census Bureau cannot give out information that identifies you or your household. The numbers we publish will not contain names or addresses.</u> Federal law protects the privacy of your answers and prevents anyone from using them against you. The back of this letter contains more information about protecting your data.
Letter B:	Your answers are confidential. A federal law provides strong protection for your census information. Your census form remains confidential for 72 years. <u>Your survey answers will only be used to produce statistics, and for no other purpose.</u> Protecting your privacy is our highest priority. The back of this letter contains more information about protecting your data.
Letter C:	Your answers are confidential, by law (Title 13, Sections 9 and 214). <u>That means the Census Bureau cannot give out information that identifies you or your household.</u> To improve census results, we might combine information from other government agencies with your answers. The additional information we receive is legally protected, just like your census answers.

**Figure 2. Versions of the Confidentiality Paragraph on the Letter’s Back<sup>3</sup>**

Letter A:	<b>Your Answers Are Protected By Law:</b> Strong federal laws require the Census Bureau to keep your responses on this form confidential (Title 13, Sections 9 and 214). This means we will not release answers that could identify you or your household. Census information must be kept confidential for 72 years. <u>Other government agencies may give us additional information about your household. We might combine this information with your answers to improve census results.</u> The same legal protections apply to any information we receive from other agencies. The combined data cannot be used to affect your rights and benefits. Any Census employee who releases your information can be imprisoned up to 5 years or fined up to \$250,000, or both.
Letter B:	<b>Your Answers Are Confidential:</b> Federal law protects your privacy (Title 13, Sections 9 and 214). These laws prevent anyone from using your answers against you. The answers you give on the census form cannot be obtained by law enforcement, by immigration, or by tax collection agencies. Your answers cannot be used in court. They cannot be obtained with a Freedom of Information Act (FOIA) request.  <u>To improve census results, other government agencies may give us additional information about your household.</u> The additional information we receive is legally protected, just like your census answers. This information is ONLY seen by people who are sworn for life to protect it.  Please visit our website at <a href="http://www.census.privacy.gov">http://www.census.privacy.gov</a> to learn more about our privacy policy and data protection.
Letter C:	(intentionally left blank—see excerpt above for the front of Letter C)

<sup>2</sup> Statements appear on the front of the letter and are underlined to facilitate comparisons across the letters.

<sup>3</sup> Statements appear on the back of Letters A and B; Letter C’s statement appears on the front. Passages are underlined here to facilitate comparisons across the letters.

## B. Summary of Findings

### General Perceptions of “Data Stewardship” at the Census Bureau

1. Most respondents expressed a general belief that the Census Bureau was trying to “do the right thing” by keeping detailed data on individuals out of the public sphere, but felt nothing was foolproof. Concern over potential, however inadvertent, data “leaks” was evident.

### Preferences for Main Confidentiality Messages

2. Response was positive for the statistical purposes statement appearing in Letter B. Preferences emerged for the “statistical purposes” statement in this letter because respondents found it most clearly conveyed that the data would only be used to compare “groups” or “averages” for particular variables (i.e., “Your survey answers will only be used to produce statistics, and for no other purpose.”). Respondents adequately and fairly consistently interpreted “statistics” in this context.
3. Letters A and C contained a statement that was also positively received: “Your answers are confidential. That means the Census Bureau cannot give out information that identifies you or your household.” This was interpreted as a definition of “confidentiality”— a concept mentioned in all three letters—and tended to clarify the meaning of this concept within the context of the survey letter.

### Preferences for Administrative Records Use Messages

4. The majority of respondents preferred letters (either A or B) with more information located on the back when comparisons were made to Letter C, which only included one brief sentence on the letter’s front about administrative records use.
5. The major issue with communicating this message was its effect on the main confidentiality pledge; respondents perceived data linking activities as an automatic breach of confidentiality because they could not imagine a one-way data capture process (from another agency to the Census Bureau). Instead, they assumed the Census Bureau provided identifying information (i.e., names, addresses, dates of birth, and Social Security numbers) to other agencies in order to “pull” and deliver the appropriate data. Though respondents felt this activity inherently conflicted with the confidentiality assurance in the letter, many already believed that government agencies communicate with one another.
6. The location of the administrative records use explanation seemed to work well in Letters A and B, where the message was located on the back. Many wished to read more detailed explanations found there, and the location of the message was not interpreted as the Census Bureau’s attempt to hide the “fine print.” Locating this type of discussion—which tended to raise more concern than any other topic in the letter—on the back of the letter allowed respondents to process the main confidentiality message first. A slight preference for Letter B emerged because it listed agencies that were not allowed to access the data (though no letter listed agencies that may give data to the Census Bureau).

## C. Recommendations

Please see the Recommendations section at the end of this report for a full list of proposed changes. An abbreviated list appears below:

1. Whichever (survey cover) letter version is selected, a field test is highly recommended before including these messages in the 2010 decennial effort.<sup>4</sup> Such a test, with a sample large enough to detect significant differences, could reveal response effects generated by negative (or positive) reactions to these messages.
2. Use versions of these messages that most closely embody plain language principles.
3. Omit unnecessary text that tends to generate extreme (or “high”) negative reactions.
4. The back of the cover letter is an appropriate location for discussions of more complex concepts.
5. The administrative records use message should be located on the back of the survey cover letter. Use the sentence pretested in the second round of cognitive interviewing on the front of the letter to alert respondents to the location of this message (i.e., “The back of this letter contains more information about protecting your data.”).
6. The administrative records use message should not be included on the census form, since there is inadequate space on the back of the form (or anywhere else) to replicate the full explanation recommended for the back of the letter. Based on an earlier recommendation from Dr. Singer, respondent materials should avoid describing the same message in various ways across the materials to keep from confusing respondents about the intended meaning of any particular message.
7. The messages from Letter B are recommended, with the following modifications:
  - a. The definitional statement of confidentiality from Letter A should follow the first sentence (“Your answers are confidential. That means the Census Bureau cannot give out information that identifies you or your household.”).
  - b. Remove statements, concepts, or words that tended to receive “high negative” reactions from respondents that were interpreted as potentially frightening, disingenuous (i.e., “over the top”), or were difficult for respondents to interpret correctly/consistently:
    - “Your census form remains confidential for 72 years.” (on front and back of letters)
    - “immigration agencies” (back of letter)
    - “...only seen by people who are sworn for life to protect it” (on back of letter and reminder post card)

---

<sup>4</sup> This is the recommendation of Dr. Eleanor Singer, consultant to this project, and the Statistical Research Division’s staff who carried out the research.



Acceptance of the proposed recommendations would result in a decennial cover letter that incorporates messages that performed well from the various letter versions we pretested. Below are excerpts of the proposed paragraphs (See Attachment I for the full letter at the end of this report).

FRONT OF LETTER (confidentiality related messages):

Your answers are confidential. That means the Census Bureau cannot give out information that identifies you or your household. Your survey answers will only be used to produce statistics, and for no other purpose. The back of this letter contains more information about protecting your data.

BACK OF LETTER (administrative records use messages):

**Your Answers Are Confidential:**

Federal law protects your privacy and keeps your answers confidential (Title 13, Sections 9 and 214). This law prevents anyone from using your answers against you. The answers you give on the census form cannot be obtained by law enforcement or tax collection agencies. Your answers cannot be used in court. They cannot be obtained with a Freedom of Information Act (FOIA) request.

To improve census results, other government agencies may give us additional information about your household. The additional information we receive is legally protected under Title 13, just like your census answers.

Please visit our website at <http://www.census.privacy.gov> to learn more about our privacy policy and data protection.

## II. INTRODUCTION

At the request of the Decennial Management Division (DMD), staff from the Statistical Research Division (SRD) designed and led a research effort to develop and cognitively pretest revised confidentiality language included in decennial materials mailed to respondents.<sup>1</sup> In revising the language, the goal was to develop confidentiality messages that communicated more clearly the nature and extent of Title 13 protections. The intent was to incorporate the new language into various pieces of the “mail-out-mail-back” (MOMB) package for the 2008 Census Dress Rehearsal, and ultimately, for the decennial census. The Census Bureau is legally or ethically required (either by way of federal law or Census Bureau policy) to convey these messages to respondents at the time of the survey request.

The DMD requested this research be carried out with the guidance of an interdivisional committee. This approach diverges from the typical process used to create content for respondent letters, which often results in a variety of stakeholders across the Census Bureau editing and approving these letters in isolation. Employing a committee to develop the confidentiality-related messages gave stakeholders, subject matter experts, and researchers the opportunity interact with one another and reach consensus on the best language and messaging strategy in light of various constraints and considerations. The interdivisional working group was assembled in the summer of 2006 to assist in the development of the new language and was called the “Confidentiality/Privacy Survey Letter Messages Working Group.”<sup>2</sup> Individuals participating in this working group possessed expertise in legal issues, policy, public relations/communication, administrative records programs, decennial operations, and survey methodology.

The desired result from this project was to generate a revised set of “respondent-friendly” confidentiality statements that: 1) the average respondent could easily understand; 2) could be inserted into pieces of the MOMB package, for instance, the survey cover letter; and 3) met the needs of the Census Bureau’s notification procedures. In particular, the working group was responsible for developing alternative messages that describe the exact nature of the data protection offered by Title 13 and the boundaries of data use. The two specific concepts used to convey these messages are that the data will be used for “statistical purposes” and administrative records may be used to supplement census data.

Toward this goal, SRD staff led and moderated meetings with the working group members throughout the summer and fall of 2006 to develop new approaches to these types of messages and consult on versions of the confidentiality messages the staff from SRD drafted and proposed. The Census Bureau’s Senior Survey Methodologist and an external consultant aided this effort

---

<sup>1</sup> The Decennial Management Division (DMD) (Kathleen Styles) sponsored this research.

<sup>2</sup> Members of the Confidentiality/Privacy Survey Letter Messages Working Group are as follows: Deborah Wagner (DID); Gerald Gates (DIR); Kimberly Higginbotham, Jane Ingold, Kathleen Styles and Nancy Sweet (DMD); Jason Machowski and Jennifer Tancreto (DSSD); Wendy Alvey, Patrick Heelen and Jenifer Santiago (OAES-POL)-- Ms. Alvey represented the Privacy Office and Mr. Heelen and Ms. Santiago represented the Legal Office; Michele Lowe (PIO); and Eleanor Gerber and Ashley Landreth (SRD). .

with expertise in privacy research.<sup>3</sup> Together, this group of individuals shaped the content and presentation of these messages embedded in the respondent materials, and the Working Group approved them prior to the cognitive pretesting phase of the project.

Three versions, or approaches, to this messaging in the letters were developed. Two letters used the front and back of the survey cover letter to communicate the full suite of confidentiality- and privacy-related information across several paragraphs, while a third letter streamlined the messaging to one paragraph located on the front of the letter. Using the front and back of the survey letter was a purposeful design choice, driven by a communicative strategy and space limitations. First, the letters were organized such that the overarching (or basic) message was presented on the front of the letter where respondents would likely encounter an initial confidentiality message. In addition, the confidentiality paragraph was limited to five or six sentences, due to space available on the front of the letter. Second, the back of the letter was used to communicate more technical or complex messages that required a greater number of well-developed statements to convey and properly contextualize the message in an appropriate fashion—thus, the back of the letter offered the volume of space required for such a message. In the case of the one-sided letter, a truncated message containing the more complex concepts was used in the confidentiality paragraph located on the front of the letter, and it was placed at the end of the paragraph to achieve a similar “presentation order” for this suite of messages.

Recreating the confidentiality message also involved reducing the required concepts to their basic components and communicating them in “plain language.” The letters also included motivational messages crafted previously by the Public Information Office (PIO).

Once several approaches to including these messages in the MOMB materials were generated, SRD staff conducted 50 cognitive interviews with volunteer laboratory respondents in the fall of 2006. The main research goal for the cognitive interview pretesting study was to ensure respondents could reliably and consistently interpret the intended meaning of various messages. In addition, we observed respondents’ reactions to individual messages and their location within the MOMB materials.

The next section of this report provides background information for the report. It discusses the scope, approaches to conveying the revised messages, and relevant research findings that were used to guide the construction of the messages. The remaining sections of the report cover the method we used to carry out the research, the findings, and recommendations for final letter wording. The report concludes by reviewing the main findings regarding our research questions, adds a cautionary note on generalizing these results to other data collection contexts, and recommends future research.

---

<sup>3</sup> Elizabeth Martin (DIR), serving as an internal consultant on the project, reviewed the working group’s messaging proposals and wrote particular passages for the letters. Eleanor Singer, an external consultant, guided the working group regarding experimental results with confidentiality assurances, resulting response rates, and implications for confidentiality messaging within the decennial context. In addition she provided feedback on the wording and placement of particular messages for the current project.

### III. BACKGROUND

#### A. Scope

The scope of the research included: 1) developing new confidentiality-related messages and applying them to mail materials used for the decennial; 2) cognitively pretesting one or more versions of the messaging; and 3) allowing the research results to inform the final version of the messages.

The Working Group was given the task of developing new confidentiality-related messages, using plain language, for the 2008 decennial mailing package, namely, the main confidentiality statement and supporting concepts. The supporting concepts either modify or describe how the data will be treated and/or used (i.e., to produce statistics and/or to study and improve process-oriented production efficiency), and how the data may be augmented by information housed within other federal agencies—concepts commonly referred to “statistical purposes” and “administrative records use” within the Census Bureau.

The Working Group also reviewed the decennial mailing pieces to determine which materials were suitable for carrying such messages. Materials under review included the advance letter, initial survey cover letter, initial questionnaire, reminder follow-up postcard, and the replacement survey cover letter, and replacement questionnaire. Because the 2008 Census Dress Rehearsal’s mailing strategy would be identical to that used for the 2005 National Census Test, the Working Group reviewed each component of the 2005 mailing package strategy for content and purpose, in addition to space available for messaging.

Once the review was complete, it was decided the initial survey cover letter, questionnaire, and the reminder follow-up postcard might be the best vehicles for the information in question. The initial mailing was selected for study because this is the point at which all of the relevant confidentiality messaging must be conveyed; the initial mailing contains a survey cover letter and questionnaire, and it is the first time respondents are given an opportunity to cooperate with the decennial effort. The reminder follow-up postcard was included in the study because it contained a pre-existing confidentiality-related message and the Working Group wished to determine how those messages performed.

The advance letter and replacement mailing were excluded from this study. The advance letter was deemed inappropriate for confidentiality language because its main purpose is to notify individuals a questionnaire will follow in the mail and provide cursory information about the decennial effort. The Working Group also felt the replacement mailing package could be excluded from the study, since it contains the same mailing pieces as the initial mailing, and the group was likely to recommended the identical messaging strategy for both mailings.

## **B. New Approaches to Conveying the Messages**

Once the Working Group identified the mailing pieces appropriate for confidentiality-related messages we set about deconstructing the main concepts and proposing and revising the resulting approaches. This task involved reviewing previous research regarding these concepts to avoid previously-encountered interpretive difficulties (see Section C below for a discussion); making certain the proposed messages were consistent with legal and policy requirements; and soliciting the input of group members with expertise in administrative data programs and communications.

The Working Group's efforts resulted in three versions of the survey cover letter. Our initial approach, informed by previous research with confidentiality-related messages, involved developing a simplified and straightforward message for the front of the letter that would communicate one main concept. More complicated concepts were located on the back of the letter so that more text could be devoted to developing these ideas and where the messages were less likely to interfere with respondents' initial impressions and interpretations of the main confidentiality message. Because the back of the letter provided more space, the confidentiality message could provide enough detail to more fully support the notion of Title 13 protections. Thus, the administrative records use message, along with a more detailed explanation of the confidentiality pledge (complete with Title 13 citations), was located on the back of two versions of the letter.

A third version of the letter was drafted in which the front of the letter communicated all the essential concepts in one short, and less detailed, paragraph. The back of this letter carried no messaging. This approach addresses a concern some Working Group members expressed that locating the administrative records use message on the back of the letter might cause some respondents to conclude the Census Bureau was attempting to "hide" information. In reviewing other locations for this particular message, the Working Group determined the back of the census form was also a candidate for a version of this streamlined messages; thus, a message was included on the form itself for pretesting purposes.

The final piece of mailing material included in the cognitive pretesting was the reminder follow-up postcard. A version of the postcard from the 2005 Census Test contained previously existing confidentiality-related messages and included a short confidentiality statement, references to Title 13, and information about Census Bureau employees having taken an oath to protect the collected data. The Working Group found these messages substantively interesting, so it was checked for accuracy and used in the cognitive interviews.

Figures 1 and 2, below, contain excerpts from the front and back of the three versions of the letter. Messages on the front of the letters contained the "main" confidentiality statement and were accompanied by various supporting statements, which included some version of the "statistical purposes" concept (portions underlined). Messages on the back of the letters were devoted to expanding some version of the confidentiality statement, citing the appropriate Federal law, and an explanation of the administrative records use concept. Letters A and B contain the front/back approach, while Letter C includes the simplified one-paragraph version.

**Figure 1. Versions of the Confidentiality Paragraph on the Letter’s Front<sup>4</sup>**

Letter A:	Your answers are confidential. <u>That means the Census Bureau cannot give out information that identifies you or your household. The numbers we publish will not contain names or addresses.</u> Federal law protects the privacy of your answers and prevents anyone from using them against you. The back of this letter contains more information about protecting your data.
Letter B:	Your answers are confidential. A federal law provides strong protection for your census information. Your census form remains confidential for 72 years. <u>Your survey answers will only be used to produce statistics, and for no other purpose.</u> Protecting your privacy is our highest priority. The back of this letter contains more information about protecting your data.
Letter C:	Your answers are confidential, by law (Title 13, Sections 9 and 214). <u>That means the Census Bureau cannot give out information that identifies you or your household.</u> To improve census results, we might combine information from other government agencies with your answers. The additional information we receive is legally protected, just like your census answers.

**Figure 2. Versions of the Confidentiality Paragraph on the Letter’s Back<sup>5</sup>**

Letter A:	<b>Your Answers Are Protected By Law:</b> Strong federal laws require the Census Bureau to keep your responses on this form confidential (Title 13, Sections 9 and 214). This means we will not release answers that could identify you or your household. Census information must be kept confidential for 72 years. <u>Other government agencies may give us additional information about your household. We might combine this information with your answers to improve census results.</u> The same legal protections apply to any information we receive from other agencies. The combined data cannot be used to affect your rights and benefits. Any Census employee who releases your information can be imprisoned up to 5 years or fined up to \$250,000, or both.
Letter B:	<b>Your Answers Are Confidential:</b> Federal law protects your privacy (Title 13, Sections 9 and 214). These laws prevent anyone from using your answers against you. The answers you give on the census form cannot be obtained by law enforcement, by immigration, or by tax collection agencies. Your answers cannot be used in court. They cannot be obtained with a Freedom of Information Act (FOIA) request.  <u>To improve census results, other government agencies may give us additional information about your household.</u> The additional information we receive is legally protected, just like your census answers. This information is ONLY seen by people who are sworn for life to protect it.  Please visit our website at <a href="http://www.census.privacy.gov">http://www.census.privacy.gov</a> to learn more about our privacy policy and data protection.
Letter C:	(intentionally left blank—see excerpt above for the front of Letter C)

<sup>4</sup> Statements appear on the front of the letter and are underlined to facilitate comparisons across the letters.

<sup>5</sup> Statements appear on the back of Letters A and B; Letter C’s statement appears on the front. Passages are underlined to facilitate comparisons across the letters.

Letters A and B in Figure 1 contain a statement that was added after the first iteration of cognitive interviewing was completed. This statement appears at the end of the confidentiality paragraph, and was inserted to assist readers in finding the text on the reverse side of the letter (i.e., “The back of this letter contains more information about protecting your data.”).

Figures 3 and 4 contain a simplified administrative records statement on the census form and a confidentiality statement on the reminder follow-up postcard, respectively.

### **Figure 3. Simplified Statement on the Census Form**

**Your Answers Are Confidential:**

Your answers are protected by Federal Law (Title 13, Sections 9 and 214). To improve census results, other government agencies may give us additional information about your household. The additional information we receive is legally protected, just like your census answers.

### **Figure 4. Reminder Postcard Statement**

Your answers are confidential and protected by federal law (Title 13, Sections 9 and 214). These laws prevent anyone from using your answers against you. Your information is ONLY seen by people who are sworn for life to protect it. Every employee takes an oath to protect your information.

## **C. Past Research & Conceptual Difficulties**

The approach taken to develop the new confidentiality-related messages was informed by past research done at the Census Bureau and by others. This research demonstrates that building appropriate confidentiality language is a difficult task. Messages must take into account respondents’ preexisting privacy attitudes, the way in which confidentiality assurances are described and/or limited, the tone and amount of detail offered, and the language or approach taken must also account for the data collection environment in which it is embedded.

We realized this current research project was unique and important for understanding how respondents—amid the backdrop of their own privacy beliefs—would receive these messages if their sensitivity toward this topic had heightened in a post-9/11 environment. The effect of privacy beliefs and attitudes has already been studied in connection with the decennial census, and evidence suggests that privacy and confidentiality concerns have an effect on respondent behaviors. In 1990 and 2000, trust in the Census Bureau’s assurance of confidentiality predicted census mail response (Fay, Bates, and Moore, 1991; Singer et al., 2001). Martin (2001) found that respondents who had concerns about privacy were less likely to fill out a Census 2000 form completely and mail it back. Similarly, in Census 2000, Singer et al. (2001) found that four factors reliably predicted nonresponse: high privacy concerns, negative views on the Census Bureau’s confidentiality practices, disapproval of data sharing, and a lack of willingness to

provide Social Security Number. These findings suggest that item nonresponse (as well as unit nonresponse) may be connected with privacy and confidentiality concerns.

In addition to respondents' attitudes about privacy, it appears the extent of the confidentiality promise, or its limitations and associated risks, also affect respondents' willingness to cooperate. Singer (1978) has found that assurances of absolute confidentiality significantly decreased item nonresponse rates and resulted in better quality data for sensitive questions. The same study revealed that a confidentiality assurance paired with a qualifying statement (i.e., answers would be confidential, "except as required by law") depressed response rates. In another experimental study, a panel given the message that other agencies might have access to their data produced a lower response rate compared to a panel that received an assurance of confidentiality "in perpetuity" (National Research Council, 1979).

Research suggests that confidentiality assurances may be counterproductive in some cases, appearing to raise rather than lower suspicion (Singer, Hippler, and Schwarz, 1992). There is some evidence that suggests stronger and more detailed assurances of confidentiality, when used in connection with surveys containing nonsensitive questions, raise suspicion and concern among respondents—perhaps because they wondered if their data might fall into the wrong hands (Frey, 1986; Singer et al, 1992).

With these confidentiality-related research findings in mind, the Working Group set out to craft messages in plain language, with a logical flow, limiting the volume of text and variety of ideas, and presenting each concept in an order that would seem cohesive and intuitive to respondents. The paragraphs below describe research illustrating some of the problems encountered when attempting to convey confidentiality-related messages to respondents, in addition to some design strategies we used to lessen interpretive difficulties and negative reactions.

### ***General Structure of the Messages***

Our approach to developing the full suite of confidentiality messages involved limiting text on the front of the letter to overarching (or main) concepts, expressed in fairly simple and brief statements. More detailed, complex, or unfamiliar messages and concepts were presented later (i.e., on the back of the letter) in an effort to control the presentation order of the messages such that respondents would be exposed to basic concepts first.

Support for this strategy stems from previous research conducted with confidentiality messages. This research found respondents preferred a simplified version of the confidentiality statement to a more "bureaucratic-sounding" version (Landreth, 2001). Both messages were located in the same survey cover letter, but the simplified version addressed this issue from a respondent's perspective by describing in plain language what types of data protections were in place. The more "bureaucratic-sounding" version contained information respondents found less useful, that is, citations to particular passages of Title 13, the phrase "statistical purposes," and unrelated messages regarding the authority under which the survey was conducted. Gerber (2003) also found the "statistical purposes" phrase by itself did not cue respondents to the concept of confidentiality, and therefore they were unable to make the logical connection between statistical data uses and the idea that their data would be protected.



From these findings we concluded that the first “confidentiality” paragraph readers encounter should be written very simply and limit content to those messages the reader can easily associate with the confidentiality concept. In addition, we wished to limit direct references to “Title 13” to the back of the letter and omit them in the main confidentiality message on the front of the letter; its inclusion in the main confidentiality message elicited mixed results. In the best case scenario, some respondents find its inclusion is reassuring, because they surmise it must be a law, though they report not knowing the specific contents of Title 13. For others, the message is less successful. Respondents have difficulty understanding how such a law could benefit them or protect them personally, and for other respondents it gives the impression the law is referenced only to protect the Census Bureau—rather than respondents. Both types of interpretations often increase the level of suspicion.

### ***“Statistical Purposes”***

Prior cognitive pretesting indicated the phrase “statistical purposes” was often misunderstood by cognitive laboratory respondents (Gerber, 2002; Landreth, 2001). Furthermore, in cases where some respondents had some idea what “statistics” were, the concept did not call to mind any association with confidentiality—which was its purpose in this context (Gerber, 2002). The phrase “statistical purposes” is complex and technical, and may raise the question of what a ‘statistical purpose’ really is. So, there is reason to believe there may be more direct and understandable ways of communicating this information to respondents.

#### Design Approaches For Pretesting Letters:

- In Version A, the phrase “statistical purposes” was omitted in favor of an approach that tries to convey the meaning implied by “statistical purposes” in plain language (i.e., “That means the Census Bureau cannot give out information that identifies you or your household. The number we publish will not contain names or addresses.”). This version focuses on the fact that only numbers will be published, without revealing names and addresses.
- While data items other than name and address can be used for re-identification after the data’s release, thereby compromising confidentiality (for example, data mining can be carried out in ways that could identify an individual by a unique characteristic like race or age), the Working Group felt conveying this concept was too complex a task and was out of scope within this context. So, additional plain language interpretations of “statistical purposes” were not explored.
- Version B substitutes “will only be used to produce statistics, and for no other purpose” for the more technical “statistical purposes.” There is some evidence to suggest people might have a better chance of consistently and correctly interpreting “statistics” than “statistical.” Previous cognitive research revealed the term “statistical” was not well understood. A principle of “plain language” asserts the use of root words facilitates comprehension, which is evidenced by previous cognitive research with another often-used term: the term “confidential” is easily understood while “confidentiality” is problematic. We also believe the phrase “for no other purpose” is a clearer statement, since “statistical purpose” had been interpreted by past respondents as vague and open-ended.

### *Administrative Records Use*

Past research with survey letters and materials used in the Survey of Income and Program Participation (SIPP) suggests certain aspects of administrative records use notification statements have produced some concern among respondents. Landreth (2002) found these statements tended to raise more concerns than the text was able to address. For instance, respondents almost always reported wanting (or needing) to know which agencies the Census Bureau planned to contact for additional personal information. Many also noted that this message conflicted with the main confidentiality message, because respondents may have assumed the two-way data sharing must occur when records are combined. It is not clear, however, how respondents will react to this message in the upcoming decennial environment. Our a priori hypothesis is that people may have a greater sensitivity about this type of government activity and potential personal consequences in a post-9/11 environment, but this certainly was not the case in Census 2000. An experimental panel contained a cover letter with an administrative records use message that did not produce significantly lower response rates compared to the control panel (Guarino, Hill, and Woltman, 2001).

#### Design Solutions For Pretesting Letters:

- Given these concerns, text was developed for cognitive testing. The design of letter versions A and B sought to minimize respondents' concerns as much as possible by contextualizing the message with strong and detailed confidentiality assurances on the back of the letter. It was not possible to use the same strategy in Letter C because this version aimed to communicate all the required messages on the front of the letter, and there were space limitations in that location.
- Letters A and B carried the main confidentiality message on the front while a secondary confidentiality message and the administrative records use paragraph were located on the back. By separating the more reassuring confidentiality messages from the more complex ones (i.e., by locating the administrative records notification on the back of the letter) we hoped to be able to control the order in which these messages were presented to respondents. We wished to expose respondents to the main message first, saving more complex messaging for later. In addition, locating the more complex messages on the back allowed us the extra space needed to fully develop the messages located there.
- Furthermore, Version A sets about communicating the confidentiality assurance by stressing the legal protections and conveying that any data the Census Bureau receives cannot be used to affect respondents' "rights and benefits."
- Version B also stresses the legal protections, but includes statements indicating the types of agencies that do not have access to the combined data. It replaces the "rights and benefits" statement with a more general statement that "the data cannot be used against you." Furthermore, we've included a statement indicating more information is available on a website, for more curious respondents.

#### **D. Limitations of the Project, the Messages, and Letter Real Estate**

The goal of this project was limited to communicating and cognitively pretesting plain language versions of a set of interrelated concepts (see below). This artifact of the research is discussed here as a limitation of the messages only because the project's requirements could not accommodate a streamlined message that briefly mentions confidentiality, much like the one cognitively pretested and used in Census 2000 (Kerwin, 1998). Instead, this effort was constrained by the number and type of ideas that needed to be explained to respondents, in addition to the limited space available in which to communicate them in the letter.

As the language development stage progressed, Working Group members reached decisions about particular concepts related to the main messages that should be conveyed in this environment to meet the spirit of the intended meaning of these concepts and be forthcoming about the limitations of these notifications. Members were guided by policy and legal requirements, in addition to ethical considerations, in an effort to convey complete and accurate confidentiality messages regarding the nature and extent of the data protections. Thus, the following concepts were considered critical for building a basic message that respondents' data will be kept confidential and it will be used for statistical purposes:

***Confidentiality.*** The notion that respondents' answers to the census form are confidential or "private" within the Census Bureau, and are protected by law.

***Statistical purposes.*** Neither respondents, nor their households, will be identified in the information the Census Bureau releases publicly; the data will be used to produce statistics (i.e., publish/release "the numbers"); and this statistical activity improves census results and programs.

***Administrative records use.*** Other government agencies may provide the Census Bureau with information about respondents; the Census Bureau will use this information in ways that enhance the data; and this additional information is also protected by Title 13.

***72 years.*** Respondents' answers to the census form remain confidential for a period of 72 years, after which they become public record.

The topics and volume of text the Working Group was required to draft tends to discuss the topic of confidentiality, we suspect, in greater detail than respondents might expect to see accompanying a questionnaire such as the decennial short form—especially in the case of topics like administrative records use and "72 years."

At the outset of this research, there was some concern over the structure these messages would take because we knew it would take space to develop them properly and we worried about the effect of a lengthy confidentiality explanation (or list of caveats). We were also aware of previous research suggesting respondent experienced comprehension difficulties with statistical purposes (Gerber, 2003) and had more questions than can typically be answered in an administrative records use message (Landreth, 2002).

To counterbalance this concern, the Working Group members attempted to craft these messages in the most respondent-friendly manner possible. We hoped to accomplish this by using plain language to the extent possible, attempting to limit the amount of text it took to convey the messages, organizing the messages in the most logical order, and providing enough supporting text to be useful and informative without overwhelming respondents with information.

## IV. METHOD

### A. Research Questions

We began designing this research with a series of questions in mind. Thus, we created the cognitive interview protocol to generate information from volunteer laboratory respondents to shed light on the following issues:

1. Are respondents able to interpret the main confidentiality-related concepts reliably and correctly?
2. How are each of these messages received?
3. Does any particular message generate negative reactions? Positive reactions?
4. How does the placement of the messages affect respondents' reactions?
  - a. In particular, how do respondents react to the placement of the administrative records use statement in its various locations (i.e., on the back of the letter, on the front of the letter, and on the back of the census form)?
  - b. Where do respondents prefer to see this message?
5. Which version of the letter did respondents ultimately prefer, and why?

Post hoc we discovered the protocol generated enough data to begin to answer the additional questions below:

6. How do these respondents feel about the Census Bureau in general?
7. Do respondents report that they trust the Census Bureau to keep their information confidential and protect their data?

### B. Pretesting Research Design

#### *Cognitive Interviews*

The cognitive interviewing method was used to pretest the three versions of the letters, the follow-up postcard, and the message embedded on the back of the census form. We originally planned to conduct the interviews iteratively (i.e., 2 rounds of 20—40 cognitive interviews total) to allow for revisions to the language based on early and prominent findings.<sup>6</sup> Once the project began, in response to a research need at the request of the Hawaiian constituency in the Race and Ethnicity Advisory Committee (REAC), an additional 10 interviews were conducted in Hawaii. Although these additional interviews were not part of the original research plan, the results proved useful for comparing findings from other parts of the United States. Staff in the SRD conducted the 50 cognitive interviews between August and November of 2006. Interviews were conducted with individuals from whom we obtained signed consent forms, and each interview lasted between 45 minutes and one and a half hours.

---

<sup>6</sup> The authors of this report conducted the cognitive interviews.

### ***Protocol***

Respondents were told they would be asked to read and provide their impressions of a couple of letters the Census Bureau was planning to mail to households, related to the census.

Respondents were reminded this research sought to understand their reactions to the content of the letters; thus, there were no right or wrong answers. To provide more background and context in which to place letters such as these, respondents were shown the decennial questionnaire as interviewers briefly reviewed the form and the types of data it collects (i.e., basic demographic information such as number of people in the household, names, birth dates, race/ethnicity data, and the like). Creating the context for which the letters, and in particular, the confidentiality messages would be interpreted was a critical aspect of the research design. We assumed the level of risk associated with providing data requested on the census form would factor into their perceptions and interpretations of the confidentiality-related messages.<sup>7</sup>

Respondents were given two letters, one at a time, which they were asked to read (silently or aloud, depending upon the respondent's preference) without interference from interviewers. The scripted cognitive interview protocol developed for this research used three methods: 1) independent recall tasks assessed topics or content respondents found salient immediately after reading a letter; 2) comparison tasks identified which messages, or which portions of similar messages, respondents preferred (and the reason for those preferences); and 3) a "feeling thermometer" task was embedded after respondents read each of two letters, which was used to obtain a general measure for the "believability" of the confidentiality messages and to observe whether content from the second letter did a better or worse job of conveying this content. Respondents were also shown the reminder postcard and the message located on the back of the census form. Scripted probes (i.e., both retrospective and debriefing questions) and unscripted (or impromptu) probes were also used to elicit verbal reports from respondents on a number of topics relating to the letters, postcard, and census questionnaire. These tools were used to generate information regarding topic saliency, interpretations of main concepts, and positive or negative reactions to particular messages and the location of these messages.

### ***Materials***

Respondents received one pair of letters during the interviews, in addition to a reminder post card and the census form containing the confidentiality-related message on the back. The three cognitive interviewers (authors) implemented a randomized interview design (See Table 1 on the next page) by administering protocol "packets" that were crafted to vary the order of each possible letter pair.

---

<sup>7</sup> We did not ask respondents to fill out the census form, rather, we wanted to make them aware of the types of data it collected.

Table 1. Randomized Interview Design

<u>Iterations</u>	<u>Letter Pairs</u>						<b>TOTAL</b>
	<u>A/B</u>	<u>A/C</u>	<u>B/A</u>	<u>B/C</u>	<u>C/A</u>	<u>C/B</u>	
First Round (20)	4	3	3	3	3	4	20
Second Round (30)*	<u>5</u>	<u>5</u>	<u>6</u>	<u>5</u>	<u>5</u>	<u>4</u>	30
<b>TOTAL</b>	9	8	9	8	8	8	<b>50</b>

\* The additional 10 interviews, conducted in Hawaii, used the following paired letter protocols: A/B (1 interview); A/C (2 interviews); B/A (2 interviews); B/C (2 interviews); C/A (2 interviews); and C/B (1 interview).

### ***Recruiting***

We attempted to recruit respondents representing a variety of social and demographic characteristics, although we focused on educational attainment level as the main recruiting criteria. We hoped this would enrich the respondent pool with individuals who would vary in their reading skill and preexisting knowledge of particular scientific concepts that we guessed might influence comprehension of the messages we intended to pretest (e.g., “statistics”).

Respondents were recruited through a variety of means. An ad was posted in the local paper and online (i.e., craigslist.com), but because we intended to research these types of messages with respondents having attained average or lower-than-average levels of education, recruiting efforts were concentrated on individuals connected to various nonprofit community agencies. Through these types of contacts we were able to recruit respondents with high school, and less than high school, educations. Some of these agencies were actively assisting respondents complete their General Education Degree, or G.E.D.

The majority of the interviews were conducted in the immediate area; 32 respondents lived in the District of Columbia’s metropolitan area. Because we suspected DC area residents, regardless of education level, might be more politically savvy by virtue of the sheer volume of information available regarding the government, eight interviews were conducted with respondents living in the St. Louis metropolitan area for comparison. An additional 10 interviews were conducted in Hawaii.

### ***Respondent Characteristics***

Of the 40 original cognitive interviews conducted for the main study, 37 were usable.<sup>8</sup> The demographic characteristics for those respondents are described below, as are the characteristics of the Hawaii respondents.

<sup>8</sup> Three interviews were conducted with malfunctioning recording devices, thus the interview data was lost.

In the main study respondents were mainly female; we interviewed 29 females and only 8 males. They ranged in age from 85 to 20 years. The median age was 33 years. Respondents from the additional interviews for the Hawaii project were mostly female also (7 females and 3 males). They ranged from 27 to 70 years of age, with a median age of 46 years of age.

Table 2 below illustrates respondents' education characteristics. We were successful in skewing our respondent pool toward those with less, rather than more, education. About half of our respondents either earned a high school diploma or the equivalent (i.e., a G.E.D.)—23 in total. An additional 4 respondents had not completed high school. Of those who pursued higher education, 15 acquired some college, but received no college diploma, and 6 respondents earned a post-secondary degree from a university.

Table 2. Educational Attainment and Geographic Location of Respondents\*

<u>Respondents' Locations</u>	<u>EDUCATIONAL ATTAINMENT</u>					<u><b>TOTAL</b></u>
	<u>Less than H.S. Diploma</u>	<u>H.S. Diploma or GED</u>	<u>Some College</u>	<u>Bachelors Degree</u>	<u>Professional Degree</u>	
D.C. Metropolitan	3	17	6	1	1	28
St. Louis Metropolitan	0	3	5	0	0	8
Hawaii	<u>1</u>	<u>3</u>	<u>4</u>	<u>0</u>	<u>4</u>	<u>10</u>
<b>TOTAL</b>	4	23	15	1	5	<b>46*</b>

\* This data was missing for one of the original 37 respondents.

In terms of race (or ethnicity), we interviewed 18 Caucasians, 15 African-Americans, 3 Hispanics, and 1 Middle Easterner (37 total). Respondents from the additional interviews for the Hawaii project mostly considered themselves Hawaiian in addition to other ethnicities: 2 respondents were Hawaiian/Chinese; 1 was Hawaiian/Chinese/Irish; 2 were Hawaiian/Caucasian; 1 was Hawaiian/Filipino; 1 was Hawaiian/Samoan; 1 was Hawaiian/Japanese; 1 reported themselves simply as Hawaiian; and 1 reported himself as Mexican.

There were differences between the respondent pools (in the original study) interviewed in Round I versus Round II; the percentage of African-Americans interviewed in Round I was higher than in Round II (55% and 26%, respectively) compared to the percentage of Caucasians (22% and 74% in Rounds I and II, respectively), all of the Hispanic and Middle Eastern respondents were interviewed in Round I, and all of the St. Louis area respondents were interviewed in Round II. Additionally, median income for respondents from the original study



varied—it was \$50,000 and \$25,000 annually for the 32 interviews conducted in the District of Columbia and the 8 interviews conducted in the St. Louis metropolitan areas, respectively.<sup>9</sup>

---

<sup>9</sup> This information was not collected for the interviews conducted in Hawaii.

## V. FINDINGS

The following section presents the results of the cognitive interviews, which outline respondents' reactions to and interpretations of specific language the letters were mandated to communicate. Afterward, respondents' preferences for versions of the letters are discussed. Following that are discussions regarding respondents' ability to navigate to the back of the letter.

### A. Respondents' Interpretations of the Main Messages

This section describes respondents' interpretations of the main confidentiality message and other supporting and/or related statements.

#### *The Main Confidentiality Message*

The term “confidential” is familiar to respondents. In these messages, we opted to use only one form of the term, the adjective “confidential.” We did not use the term “confidentiality” because previous research had indicated that some respondents were unsure of the meaning of the word in this form. However, the term “confidential” itself was generally readable to these respondents. During these interviews, we did not discover any respondents who stumbled in reading this term, or claimed not to understand its meaning.<sup>10</sup>

Respondents use several strategies to define “confidential.” Although they sometimes said that “confidential” meant “secret” or “privileged,” the most commonly used term in this context was “private.” The Census Bureau frequently uses the terms *privacy* and *confidentiality* to mean two different things; *privacy* refers to non-intrusiveness in data collection and *confidentiality* refers to protections after the data is collected. These respondents, however, did not make this distinction:

“Information will be kept totally private – it’s privileged information.”

“Secret...not publicized.”

“Secret...not allowed out, secret, no matter what.”

“Your personal information is private.”

“My privacy is protected.”

“...there shouldn’t be any breach of your own privacy.”

Another less frequent approach to the meaning of confidentiality is that a data collection has to be anonymous to be confidential. This was encountered rarely, but for at least one respondent, the collection of name voided the promise of confidentiality:

“...it says confidentiality, confidentiality means that you don’t have to really put your name, cause you won’t know who it is...So this is not really confidentiality, because you are putting your name up there.”

---

<sup>10</sup> It should be noted that this group of respondents were recruited as literate in English; respondents just learning English or reading the term in translation might not have similar familiarity with the term.

Basic interpersonal definitions must be extended to fit a bureaucratic context. The concept we are attempting to communicate is familiar to respondents, but their ideas about it have not been formed in the context of survey data. This is important to our messaging because respondents' original ideas about confidentiality form a starting point and background for what we are trying to communicate.

When we questioned them on their understanding of the term “confidential,” respondents tended to rely on definitions drawn from the realm of interpersonal communication. Their expectations of confidentiality appear to be based on an image of one person telling a secret to another. Thus, they say, “maintaining my information is strictly between you and me, that’s what [confidential] means,” “just between us,” or “you don’t tell it out anywhere.” This interpretation has two consequences. First, the idea of confidentiality is generally assumed to be unconditional, that is, it should apply to all persons at all times:

“People are very, very, very, very concerned about their privacy and I think they wanted to make them feel assured that the information would be made confidential under all circumstances no matter what...”

“...means it should be kept between you and I and you’re not to repeat it to anyone else under any circumstances.”

But this unconditional, two-person image poses an immediate problem for respondents. They understand that others must see and use survey data, and that they are giving information to a large organization. Thus, respondents must expand definitions of confidentiality drawn on a model of telling secrets to a trusted interlocutor, in order to apply it to the messages being transmitted in the letter. The process of expanding the definition of confidentiality is easy to see, as respondents try to assess how wide their circle of protection is going to be. The following quotes indicate the process of including a wider group of people; these respondents seem to waver between their understanding of confidentiality as applicable to ‘you and me’ and to the Census Bureau as a larger entity:

“Confidential means that it’s between you and that person and nobody else, or you and the Census Bureau and nobody else...Maintaining my information between you and me, that’s what it means.”

“Whatever you write down is gonna be between you and the CB, the person in the Census Bureau...means it should be kept between you and I and you’re not to repeat it to anyone else under any circumstances.”

Thus, defining the extent of the confidentiality promise is a cognitive problem for most respondents. They must process the idea of where the data goes and who is allowed to see it. They do not always resolve the problem in the same way. They generally expand the confidentiality idea to apply to at least some of the Census Bureau’s employees, although they are not certain if it is based on relevance, need to know, or if everyone in the Bureau will have access to their forms. The following are examples of respondents who believe that only some of the Census Bureau’s employees will have access to the data, based on need-to-know:

“Only the people that this survey or census form is important to.”

“That it’s private, it’s only going to be discussed or disclosed with the people that it’s relevant to, it’s not going to be disclosed to anybody else that doesn’t need this information.”

However, most commonly, the resolution of this problem is that the information will remain within the Census Bureau taken as a whole:

“It’s kept within the Census Bureau, it goes nowhere else.”

“Pretty much nobody but the Census Bureau can get this information.”

“[Identifying data] won’t be given to the public—[the data] will be within the Census Bureau itself.”

This conclusion is not entirely certain for some respondents. They are not entirely sure if their conclusions about who is included in the promise of confidentiality are correct. The uncertainty is evident in the following:

“Confidential—like no one outside of the Census Bureau will read it?”

[Interviewer Probe: “Who can see your answers?”] “Um, the Census Bureau. Anyone who works for the Census Bureau and other governmental [sic] that are connected to the Census Bureau... Well, I’m not really sure who it is. I don’t know much about who is connected to the Census Bureau.”

This uncertain scope is important, because it bears on respondents’ subsequent reactions to the ideas about data sharing within the government. Their natural definitions of confidentiality must be expanded to fit a bureaucratic environment; and the definition of “confidential within the Census Bureau” is not automatic or well established.

### ***The “Statistical Purposes” Concept***

The letters included some ways of supporting the promise, or idea, of confidentiality given on the front of the letter. Letter A included a sentence that served as a definition of confidentiality: “That means the Census Bureau cannot give out information that identifies you or your household. The numbers we publish will not contain names or addresses.” Letter B included the phrase: “Your answers will be used to produce statistics, and for no other purpose.”

These supports for the idea of confidentiality were intended as replacements for the phrase “used only for statistical purposes,” which has been used in prior decennial and survey letters. This phrase was intended as a way of telling respondents: 1) that identifying data will not be released; and 2) that their data will be used analytically in a variety of contexts and not just Census tabulations (e.g., to improve processes and other statistical programs). Previous research had indicated that the term “statistical purposes” was difficult for respondents to understand (Gerber, 2003; Landreth, 2001). Therefore, these statements were revised to include the “plain language” expressions of these concepts, quoted above.

“Used to Produce Statistics and for No Other Purpose.” Respondents did not experience confusion about this phrase, and generally thought they knew what it meant. Their explanations of the meaning tended to be technically inexact but they were able to provide definitions that were generally in the right area (at least if the ‘ballpark’ is generously defined). They responded in terms of “numbers” and “adding things up” or more vaguely, categorizing things. Respondents sometimes paraphrased this concept by saying:

“Statistics means numbers and that really all they’re looking for are just the numbers.”  
“...It’s like, putting things in buckets.”

They often responded to this phrase in terms of the uses of the data that they anticipated, based on the questions they had seen in the decennial questionnaire we had just shown them. People responded in terms like, “who lives, who dies” or imagined situations in which a school district needed to know how old children are, or what kind of ethnicities are represented in a particular community. This phrase is useful, because it appears to connect respondents directly to the idea of the ultimate uses of their data. It does not, however, appear to carry the connotation of removing identifiers from a large data set.

It is not clear why the phrase “used to produce statistics” worked so much better than “statistical purposes.” This may be a question of grammatical form. Just as “confidential” is more familiar than “confidentiality,” the term “statistics” may be more familiar than the adjective “statistical.” In addition, the context of the term “statistics” may enhance its understandability. It seems likely that the phrase “statistical purposes” was confusing to respondents primarily because they could not understand what a *statistical* purpose would be, as opposed to any other sort of purpose. When we discussed the second part of the phrase (“and for no other purpose”) with the current respondents they sometimes expressed brief confusion. What else could you do with data but to count it up? In the current formulation, respondents easily dismiss this concern, however, because those *other* purposes have been clearly delineated as events that are not going to happen.

It should be noted that very few respondents actually used the term “statistical purposes” during our subsequent discussions with them. These respondents seemed to have greater sophistication in their understanding of data processing than our average respondent. In these few cases, it did not appear to carry a different meaning than the phrase included in the letter.

“That means the Census Bureau cannot give out information that identifies you or your household.” This sentence in Letter A proved very popular among respondents. This was seen as a definition of what we meant by “confidentiality.” Many people remarked that this statement was helpful for communicating an exact definition of what the Census Bureau means by confidentiality in this context. Since, as we have seen, they must process how this promise will be enacted in a large agency, having this definition at the outset was very useful. It clearly conveyed to them that names, addresses, and other identifiers such as phone numbers and social security numbers<sup>11</sup> would not be divulged. Since the respondents have to process the idea of

---

<sup>11</sup> In fact we do not ask for social security number, and respondents all had an opportunity to examine the questionnaire before we began the interview. Many, however, came to believe during the interview that they must

confidentiality to understand what a big government agency might mean by it, it was useful to have this definition.

This statement was also highly salient to respondents. When asked to compare versions of the letters, those who chose Letter A often pointed to this sentence as the element that had swayed their choice. In addition, those who chose other versions sometimes remarked that we should include this sentence in their favored text.

The second sentence in this “definition” was “The numbers we publish will not contain names or addresses.” This did not appear as useful as the sentence that preceded it. Respondents were confused about what the “numbers” were, and it appears that this term is more vague than “statistics.” Respondents also indicated that they did not know what the term “publish” meant, under these circumstances.

Thus, the statement appearing in Letter A was successful in providing support for the idea of confidentiality by clearly communicating that confidentiality precludes the release of identifying data. It is less successful, however, in providing respondents with the information that their data will be used in analysis. (Both of these elements were intended by the phrase “statistical purposes”.) The statement in Letter B “used to produce statistics and for no other purpose” was more successful in this, since it brought to mind a wide variety of data uses. As a result, we believe that both the “definition” provided in Letter A and the “statistics” explanation in Letter B are necessary to convey the entire concept for respondents. It should be noted, however, that respondents did not understand any of these phrases to include the use of data beyond what is collected in the current survey. There was no evidence that they interpreted these communications to mean that data from other agencies or other data collections will be used along with Census data.

### ***Administrative Records Use Message***

One important aim of this research was to examine respondents’ understanding of new confidentiality-related language, which notifies respondents of our use of administrative records data from other agencies to augment the analysis of Census results.

The descriptions of administrative records use were as follows:

“Other government agencies may give us additional information about your household. We might combine this information with your answers to improve census results. The same legal protections apply to any information we receive from other agencies.” (Letter A.)

“To improve census results, other government agencies may give us additional information about your household. The additional information we receive is legally protected, just like your census answers.” (Letters B and C)

---

have seen a request for social security number. This may have been the result of the stress we were putting on confidentiality: if the data are not really sensitive, this degree of emphasis does not make sense.

The administrative records message appeared on the front of letter C (since no information was presented on the back of this letter). It was on the back of Letters A and B.

What we intended to communicate in these phrases was that the Census Bureau might receive administrative records data from other agencies. The agencies involved were not specified, nor were the specific uses to which the data would be put. The stress was on the protections for the combined data. We wanted respondents to understand that data flowed only one way, and all such data received by the Census Bureau would be protected by Federal law, that is, Title 13.

Communicating the administrative records use message proved to be less straightforward than defining confidentiality. Three important issues were raised by this notification. First, we were concerned with the understandability of the messages. Did respondents get a clear picture of what we intended to do with the data, and of the protections that we afford to it? Second, we were concerned with the acceptability of the message. In particular, did it affect respondents' understandings of our promise of confidentiality? Third, did the placement of the letter affect the understanding and acceptability of the administrative records notification? We were not certain how the placement of the message on the back of the letter would be received by respondents.

Respondents' understanding of the administrative records use messages. One consistent finding was that this message was vague; respondents wanted more information and found all the explanations lacking certain information. Their initial reaction to the sentences was often to question which agencies would provide more information to the Census Bureau. Sometimes they were genuinely puzzled by the letter's lack of clarity on this point:

“What government agencies are you talking about? You all didn't pinpoint what government agencies?...What if you have never been in any kind of government, so how would you combine their information?”

“Government to me is government. I'm not knowledgeable, but...I'm sure it's a lot of them...If you want us to get that, you have to put it on the line.”

When respondents were asked which agencies they thought might be called on to provide additional information, they mentioned a wide variety of Federal and local agencies, in addition to other sources of data. Data sources respondents mentioned included local law enforcement agencies, the FBI, the DMV, the Navy, the Veteran's Administration, Social Services, Medicare, Medicaid, “Section 8” housing, the IRS, Social Security, Child Protective Services, and employee records for anyone who worked for government. Private organizations, such as real estate companies, banks, and other creditors, were also mentioned. Although any particular respondent mentioned only one or two “agencies,” it appears that the overall impression may be of a wide search for data from many entities, both governmental and nongovernmental.

In general, once respondents learned the Census Bureau intended to collect additional information from other government agencies, they typically wanted more information about these activities rather than less. The majority of respondents preferred letters (either A or B) with more information located on the back of the letter when comparisons were made to Letter C, which included one brief sentence on the letter's front about administrative records use.

While some actually liked Letter C for its brevity, many respondents appreciated having access to a lengthier explanation located on the back of the alternative letter.

Each treatment of the administrative records use message, however, avoided naming the agencies the Census Bureau intended to contact—information respondents often sought. Though none of the letters provide these specific references, preferences for Letter B emerged because it at least mentioned some agencies that were not allowed to access the data (e.g., tax collection and law enforcement agencies).

The purpose of the administrative records use program was unclear. Respondents were also puzzled about why we might be getting data from other agencies. Because we had framed the purpose as “to improve census results” this led to the question of what the Census Bureau was trying to improve. For instance, one respondent remarked:

“First of all, I don’t know what you are trying to improve?...What are you trying to make better than the Census, say in 2000?”

The text did not provide an explanation, so respondents were left to imagine for themselves what the purpose of the program might be. For the most part, they assumed the reason for getting additional data was to check respondents’ specific answers. For some, this was connected with the idea that someone might have provided misleading data deliberately, or suppressed information:

“Maybe if you’re—I’m not going to say dishonest, but maybe if you’re—yeah, I’m just gonna say dishonest, other government agencies, the Census Bureau can go to other government agencies to get more detailed information about your household.” (IRS example)

“Even though people can’t get the information on you from the Census Bureau, the Census has rights to get your information from whoever it needs...Maybe you’re on welfare or Medicaid or Section 8 housing, maybe if you didn’t fill out the information for the Census Bureau they could probably contact the government to find out who you are...”

Thus, the major idea about the purpose of the data transfer from other agencies was that the Census Bureau intended to examine the accuracy of answers to Census questions. Most respondents believe that they would never have any reason to lie or mislead the Census Bureau—particularly after seeing the simple demographic questions included in the census form we showed them. Nevertheless, the notion that the Census Bureau would “check” their responses was uncomfortable for some. They believed that *they* were the best source for accurate information, and that errors might be produced by going to other agencies. For example, one respondent mentioned that an elderly parent had recently left her household and moved in with a sibling. She questioned whether any records might include him and whether we would inaccurately list him with her household.



The conclusion of this logic seems to be that if we are “checking” answers, the best source is the respondent him- or herself. As a result, we should be asking anything we want to know on the questionnaire:

“If you’re trying to keep it as secure and confidential as possible, you need to call the person who filled out the survey – don’t go to other places that you shouldn’t be.”

“[It] opens the doors for questions...why do you need to go get other information?...you should just get it from them, then – or you should just ask it on the form.”

It should be noted that almost all of the respondents understood the data transferred to be specific, identified data. The concept of a large data set with no identifiers was not mentioned by these respondents. It is doubtful that they had a sufficient knowledge base about data processing to create this association. In fact, the language used in the letter to explain the administrative records program has no cues to respondents as to whether it refers to specific identified data or a large data set.

Direction of the shared data. It was our intention to inform respondents that the data transfer was entirely one-way; the Census Bureau would receive, but not give out any data. In fact, there were some respondents who did read our messages this way. Even though they may have thought of the transfer of individual information, they understood that the information would remain within the Census Bureau:

“They told me that other sources could be called on, but that’s confidential too.”

“I don’t know how many government agencies there are but I’m sure there’s a lot. So it means that [the Census Bureau has] access throughout the government, not that they would do anything with that information, but they would have it available if needed.”

“There are agencies out there that they can inquire as to your identity or who you are, or, cause Lord knows, there’s enough ways to figure out where people are or who they are. And they can use our information...just to improve the census.”

However, many more respondents understood the data to flow in two directions: both to and from the Census Bureau. Thus, one respondent paraphrased the administrative records use message directly as “they might give your information to other government agencies.” Others came to the same conclusion after pondering the true meaning of those initially confusing sentences. A very common interpretation was that, if the Census Bureau received data, they must also give it out. Respondents reasoned that if the Census Bureau were to check their answers, or get new information about them from other agencies, the Census Bureau must at least be providing the other agencies with respondents’ names and addresses:

“Somebody else can give them more information about you. But then that’s sharing your information so would they know [what to tell Census Bureau] unless you guys asked them? How could they give out other information unless you guys are asking

them?...The Census Bureau is asking another part of government about you, so maybe ya'll work hand in hand."

Acceptability. The acceptability of the message was to some extent dependent on the assumptions that respondents made about the nature and conditions of the data transfer between agencies. As indicated in the previous section, respondents assumed the Census Bureau at least had to reveal name and address in order to receive the requested information from other agencies. This was concerning to some of them because it was regarded as a priori proof that the Census Bureau was lying about keeping their data confidential. This logic led some respondents to believe that this program breached the promise of confidentiality:

"Ok, if it's legally protected, that means that should no one see it, but the Census Bureau and me and you. So how can it be legally protected if you are combining agencies?...They're telling you one thing but they're doing something else."

"The same protections apply [to data received from other agencies], but the other agencies are giving it to you. So, in return [is the Census Bureau] allowed to give it to the other agencies?"

"I'm going back to where government agencies can exchange information, at will, and what—half of the entire population of the DC area works for the Federal government. So, EVERYBODY has access to my information."

Some respondents also thought that we must be providing Social Security numbers (SSN) to the other agencies in order to request data about them. It is worth noting that after thinking about this data transfer, respondents often assumed the census form collected SSN. We had taken care to make sure that respondents looked over the questions on the short form, and even recounted the topics to them, but at this point in the interview they no longer recalled that SSN was not on the form. In fact, some of the respondents looked back at the questionnaire to see if they could find it, and "confirm" this inaccurate recollection.

To some extent, the notion of two-way data sharing is rooted in a common belief that respondents have: all government agencies share data. The sharing of data among government agencies is not only assumed, it is often considered highly legitimate. Respondents often make the assumption that data are shared not just among the Federal government, but with state and local governments as well. "Government" in this view is all one thing, and of course data will be given out to those who need to use it. This is perhaps the widest extension of the confidentiality concept: "confidential with anyone in government with a need to know." Such respondents often explicitly tell us that they trust the government, either because of personal experience (such as being in the military) or as a matter of patriotism.

"The Census Bureau is asking another part of government about you, so maybe ya'll work hand in hand...if you guys work together, then I feel it's gonna be very confidential, it's gonna be protected."

“Hopefully the agencies will work better, um, I’m hoping they’ll work together because they should. If they’re government agencies they should work together.”

“I guess if you had participated in things with other agencies that it might be good to combine the results to see if they were accurate and see if there was anything noticeably different that you might have left out or that changed from the time you filled out the census.”

“I’m thinking government agencies are sharing information...that sort of means that even though you said my answers are protected, all the government agencies can share information.”

“I think [census answers] would go any way that would give the most potential information to whoever—whatever agency—in the government it would serve the most benefit to.”

The assumption of two-way data transfer destroyed the concept of confidentiality only for a minority of respondents. It had much less effect on respondents’ opinions than we might have anticipated. One reason for this was that some respondents took it as an assumed part of the way that they understood government to operate. A few respondents thought they were reading that the data exchange was only one-way; they doubted that they could have understood correctly:

“So I would assume it would be available to any government agency [but] the way I’m READING it, it says, “Census”...but it’s the other government agencies. Census works for the other government agencies, yah?”

Thus, the assumption of government data sharing is for many respondents just “business as usual,” and they sometimes read the administrative records statement in that light. These statements are seen as a necessary notification of what may be done with an individual’s data. For some, the “admission” that the Census Bureau is doing what they already assumed all of government does lent this agency certain credibility. Others thought the administrative records sentences were there primarily to protect the Census Bureau. For example, some respondents indicated that they are used to seeing such “fine print” in privacy statements sent to them by banks and credit card companies, etc. So, it is no surprise that these statements were interpreted as “the stuff you have to say” and the “fine print,” which explains the exceptions and conditions that respondents predict they must be notified of:

“Well, it’s letting you know you know that, first of all they’re telling you that the Census Bureau is going to keep your confidentiality, but they’re letting it be of no surprise to you that an employee that might not be...a census employee may view your records...”

Trust in government. For some respondents, the pledges of confidentiality were affected by preexisting ideas about the trustworthiness of government. These ideas were both positive and negative, and affected attitudes toward the confidentiality pledge in both positive and negative ways. For some respondents, a government promise increased the credibility of the assertions:

“It’s set in stone in federal. Everybody can’t use federal...since it’s federal, I would think everything would be legitimate.”

“I trust the government.”

“Well, with identity theft now, I don’t know, it just really seems like anyone [can see the information]. But my government is telling me that they’re going to, you know, protected my information, so I’m hoping, you know.”

But for other respondents, the government was not seen as trustworthy:

“Look, I’ll give that a 3 [medium rating], because how far does your trust in government go, ok? You know, things have a way of slipping out.” (in reference to recent data spills in the news)

“It’s not that I have anything to hide, but when you’re talking about the government that’s a big, huge body, and it’s scary when you think they are collecting information about you. To me it’s like having your information out there in cyberspace, I don’t necessarily believe that they will protect it.”

“[The message means] you can tell the government anything you want to, and they will not pass that information along to anyone else...some people will have a problem with that, because they just don’t believe that what the government says – that they will keep it confidential. I don’t believe it myself...[information becomes] part of the system.”

To some extent, these views represent a worry about government data sharing that respondents had maintained for some time. The themes in the statements above are typical: distrust of government promises, the belief that data in government computers are available all over, and that mistakes will lead to things “slipping out” no matter what the policies say. The last of these themes was particularly evident among our respondents. Familiarity with incidents involving data leaks was almost universal and had a clear effect on respondent trust in government promises:

“I still have doubts on certain things, because of the things that happen, like I said, identity theft and stuff like that, so I know that our information, there have been so many leaks, for different things in the government that I can’t give that a total protection. Because of things that have happened. But I think that every effort is made to protect it.”

“[The VA laptop incident proves] you can never be sure your information is 100% safe...If somebody wants to get it, they will get it.” (in reference to the Veteran’s Administration [VA] stolen laptop)

“There’s always something that happens. There’s always a crack in the system...there’s always room for error.”

The most commonly mentioned incident was the theft of the VA laptop, but respondents also frequently mentioned media coverage of other data problems, such as leaks involving bank or credit card data. Although none of the incidents they recounted involved the Census Bureau, respondents have come to believe that data security is uncertain at best. As a result, even if they accept that the Census Bureau is “trying to do the right thing” often they do not believe that we can promise complete data security.

A second element also affects trust in governmental promises. Citation of law was generally popular with respondents. However, some respondents realized that the laws involving privacy have changed, and thus, any particular law might be changed or over-ridden at any time:

“Do you think all these laws about privacy mean anything now that we have the Patriot Act?...[It] eliminated a lot of these laws.”

Respondents are familiar with some changes in the legal privacy context, such as the Health Insurance Portability and Accountability Act (HIPAA). They had experience with this as employees and as patients. Employees in the health fields have been trained about patient confidentiality. If they had positive experiences with this, it tended to reinforce our messages about protecting data. Others had less positive experiences; and it “proved” to them that employees could not be trusted to keep information private. For example, one respondent indicated that his HIPAA training had included a rule not to talk about any patient, even if that person was well known to other staff, who might be legitimately concerned about the person’s condition. This respondent had seen many cases where this rule was ignored, and did not think that Census employees would be any different. The idea that “people talk” and policy cannot stop them from doing so was also expressed by respondents with experience in other bureaucratic situations (e.g., former Federal Human Resource employees).

Salience of the message. In assessing the acceptability of the administrative records sentences, we were also concerned with the salience of the information to respondents. That is, we wanted to know how immediately striking or important the information seemed to be to them. In cognitive interviews, the one of the best measures of salience is the degree to which respondents spontaneously mention the information on first reading the text. Another measure is the degree of concern or interest they express.

In this instance, a few respondents noted the administrative records sentences right away and began discussing them before they had finished reading the rest of the text. But it is important to note that this was relatively rare behavior. Although some respondents expressed concern directly after reading the statements, more often they did not do so until they were probed about them. Their first remarks were usually about other information contained in the paragraphs. The penalties statements in Letter A and the explanation of the agencies that could not get Title 13 data in Letter B appear to have been much more salient than the administrative records statements.

During our questioning, we asked respondents if they read the sentences or skipped over them. They generally told us they had read the text, but did not pay much attention to it. (Without eye-tracking information, it is impossible to test the accuracy of this assertion, but respondents

seemed genuinely convinced that they had looked at the sentences in question.) If respondents expressed concern, it was usually after we discussed the content of those sentences for a while. It is possible that respondents must process of the statements more deeply in order to arrive at the interpretations that distress them.

There are several possible reasons that might account for the lack of salience of the administrative records use message, which perhaps all coincide. First, respondents had already processed the simple confidentiality message on the front of the letter, and concluded that the back of the letter was amplification, or “the details,” about what they had already read. Thus, they were not expecting information that might be at odds with the original message. Second, respondents expect to see the “fine print” (as described above) and are well practiced in ignoring it. Third, respondents had more positive messages on the back of the letter to process, and gave initial attention to these. Letter B was particularly effective in this regard. The statements appearing in that particular letter, about other agencies that cannot receive your data, preceded the administrative records statements directly and received a lot of attention from respondents:

“The answers you give on the census form cannot be obtained by law enforcement, by immigration, or by tax collection agencies. Your answers cannot be used in court. They cannot be obtained with a Freedom of Information Act (FOIA) request.” (Letter B)

We had the sense that in reading the back of the letter, respondents were more aware of this welcome information, and often did not focus on the administrative records statement until we pointed it out to them. It should be noted, however, that we observed Hispanic respondents who did not view a portion of this paragraph as positive or benign. They worried about the inclusion of the phrase “immigration agencies” (located in the first sentence in the paragraph above) in the statement describing which government agencies would not be able to access identifiable census data. Whether or not a person is an efficient or practiced reader of English, these respondents felt many Hispanic recipients of this letter would recognize the word “immigration.” Given the negative connotation of this term in the context of a government document, our Hispanic respondents felt the inclusion of this word would alone dissuade potential respondents, regardless of the positive context the creators of the paragraph intended to invoke.

Location of the message. The placement of these statements was important in determining respondents’ reactions to the messages. They were better received if they appeared on the back of the letter, after respondents had a chance to process the simple confidentiality message, and were part of the text that the respondents had already decided was “the details.”

“Yeah, this is supplemental to what they said [on the front of the letter].”

“I think it’s just an extension of the last sentence on the other side (Letter A). In other words, they’re trying to tell you how they’re—I said they don’t tell you how they are going to do it—but this is going into a little more detail.”

“[It’s the] same, but stronger [than text on the front].”

The location of the administrative records use explanation seemed to work well in Letters A and B, where the message was located on the back, for several reasons. Locating this type of discussion—which tended to raise more concern than any other topic in the letter—on the back of the letter allowed respondents to process certain information first: the main confidentiality message. In addition, respondents reacted positively when presented with the option of reading the back of the letter for “additional” information.

We were initially concerned that the placement of this information on the back of the letter might be interpreted as an attempt to hide important information, but respondents did not interpret it this way. In fact, we never observed negative reactions to the organization of the messages. Although respondents did think of some of the information as “the fine print” they did not conclude that the placement of the message was deliberately deceptive. This was true even when respondents were alerted to the administrative records use message on the back of the Census form. Regarding the text on the back of the letters, one respondent pointed out that the large font size (same font used on the front) communicated to her that the Census Bureau was not trying to hide anything. She noted that if it the letter truly intended the text on the back of the letter as “the fine print,” then she would expect to see a much smaller font size used for messages located there.

#### ***Federal Law: The “Title 13” Concept***

The term “federal law” was mentioned on the front and back of Letters A and B, and “law” was mentioned on the front of Letter C. Specific references to “Title 13” often accompany statements regarding confidentiality, a common practice in respondent communications from the Census Bureau, but Title 13 references were limited to the back of Letters A and B (they did not appear on the front). Letter C specifically referenced Title 13 on the front because there was no text on the back.

Respondents generally understood there were laws in place to protect them (e.g., “Saying they will not release my identity to anybody...Census Laws.”). The term “strong” was paired with the phrase “federal laws” on the back of Letter A, which gives a nod to the notion that Title 13 was a special law designed to protect respondents’ data (though this fact is never explicitly stated and respondents never infer that it is a special law unique to the Census Bureau), but the reactions were mixed. Some respondents reacted very positively to the phrase “strong federal laws,” reporting that the word “strong” did impress upon them that the law was taken seriously. Others reported that the presence of this term in this context brought to mind the contrasting idea of “weak” laws, and this was unsettling for respondents because they reasoned that there should be no difference between laws in weak or strong terms. In a few cases respondents reacted negatively to the term’s inclusion; they indicated the phrase sounded “over the top,” surpassing the limits of reality due to this small linguistic flourish.

While it was helpful for respondents to learn that laws exist to protect their information in the decennial context, they relied on surrounding text (i.e., “federal law”) to interpret the unfamiliar references to “Title 13” on the back of Letters A and B and were unable to specifically identify the contents of Title 13:

“It’s a federal law.”

“It’s part of the confidentiality protection area.”

“This is the law and I guess it lists the number where the law is cited.”

“It’s a specific part of the law...the Federal law is broken up into tons of bits and pieces and this just goes to a certain, um, piece of all that—I would say that this is regarding privacy maybe and protection of rights or somewhere along that group.”

“It looks to me like it’s a federal law, I guess part of a sentence in the law.”

Respondents were unfamiliar with Title 13 or its contents, and this finding is consistent with previous findings from other respondent letter research on this particular topic (Landreth, 2002). Some respondents guessed they could look it up at the library or on the internet, but few reported they were willing to track down the actual text of the law. The reference, by itself, provided little information and so respondents could only make general statements about the intended meaning which may have otherwise been impossible if it had not been accompanied by the phrase “federal law” or “law.”

The location of the reference to Title 13 seemed to make a difference in respondents’ interpretation of the main confidentiality message. While a few respondents liked seeing the Title 13 reference on the front of the letter, many more respondents interpreted it as legalese. The interpretation of Title 13 as a bureaucratic disclosure tended to suggest that the surrounding confidentiality messages functioned more as a “protection” for the Census Bureau rather than respondents. In this way, the Title 13 reference on the front of Letter C may have distracted respondents from the main confidentiality message, compared to those who saw the reference on the back of Letters A and B. Respondents who saw the Title 13 reference in these letters had previously been exposed to the main confidentiality message on the front of the letter, and so they tended to interpret the unfamiliar title as “the details” or “the fine print.” So, even when the title appeared on the back of the letter, it did not provide additional information that was meaningful to respondents, but at least it did not seem to interfere with their interpretation of—or serve as a distraction from—the main confidentiality message on the front of the letter.

### ***The “72 Years” Concept***

This message was included in the letters because it clarifies the extent to which census data will be kept confidential. The message intended to convey the data would be kept out of the public sphere for quite some time—72 years. This statement appeared in Letters A and B only; space issues kept this statement from being included in Letter C. Respondents varied in their perceptions regarding what happens to the data after 72 years, most likely because the statements were written fairly vaguely:

“Census information must be kept confidential for 72 years.”

“Your census form remains confidential for 72 years.”



Some respondents guessed correctly that the information would be released after this amount of time while others assumed it would be destroyed:

“After 72 years [the statistics] are available—the census form.”

“[After 72 years] I hope they throw it away.”

Regardless of their assumptions, respondents often told us they were uncertain what would happen at the conclusion of 72 years and many indicated they would be “dead by then,” and so they reckoned that revealing their personal information would not result in negative consequences

“Think about it, if you fill out a census form in your 20’s [what’s the] likelihood of you caring about it in your 90s?...so you feel pretty safe from that aspect.”

“I’ll be dead by then, so I won’t care.”

While part of the main message seemed to be getting through, respondents were often distracted by what struck them as an odd statement. They tended to spin off in different directions thinking aloud about why it was included in this letter and what it really meant. One respondent wondered why the Census Bureau would keep information that long to begin with, since she was sure personal information was outdated after seven years. Many respondents were also surprised the data were kept that long. A consistent sentiment from respondents was that “72 years” seemed to them such an arbitrary and number, that many of them spontaneously mentioned it before we had a chance to probe for their reactions:

“...72 years, where the heck did that come from?”

“Oh, and I wonder where they came up with the 72 years...How did they pick 72 years? Who came up with that?”

“Why 72 years? I know I’m not going to live...but why 72 years?”

Some respondents were so struck by the arbitrary nature of “72 years” that they found it humorous. This reaction occurred consistently enough that it led us to conclude that, as a general principle, messages should strive to avoid sounding bureaucratic and funny at the same time.

Respondents were not only distracted from the main point of this message due to the “72 years” portion of the statement, but when probed, they also could not produce any reasons why the data might be useful at the end of this period. While some respondents generally understood this statement to modify the confidentiality statement, it would have helped them to know to what use the data might be put at the end of this time (e.g., historical research).

## **B. Motivational Messages**

These messages were included in the letters and were meant to provide motivation to respondents to respond, add emphasis or credence to main messages, or may have acted as motivational text due to its perceived usefulness. Some elicited extreme reactions (both high and low) that may render them less consistently useful in the context of a survey letter, while others were quite consistently well received. In deciding whether to include messages that elicits both positive and negative reactions, the conservative decision would be to exclude unnecessary text that functions in this way.

### ***“Protecting Your Privacy is Our Highest Priority”***

This message was located on the front of Letter B. Very few respondents noticed and commented on this particular message spontaneously. The conclusion we draw from this is either that the message was not very salient to respondents, or it failed to carry much weight. This finding suggests the statement may not be a very effective message. When probed about the message directly, most people found it to be a positive one and only a few people suspected this message overstated the Census Bureau’s commitment as expressed here. In fact, these respondents snickered when they read it, which indicated they had strong negative reactions to this message. Most respondents, though, interpreted it as a commitment on the part of the Census Bureau, and some even cited the message as one that would encourage people to fill out the form.

### ***“Sworn for Life”***

This message, “Your information is ONLY seen by people who are sworn for life to protect it”—located on the reminder postcard and on the back of Letter B—generated extremely negative reactions. This was due to a misinterpretation of the intended message for some and outright disbelief for others. In the case of the former circumstance, the concept of being “sworn for life” was interpreted by some respondents as “protecting the data with your life.” Respondents had a difficult time making the logical connection between the Census Bureau’s employees’ actions or regular duties and putting their lives on the line for data (or as one respondent put it, “taking a bullet” for the data). For this reason, the misinterpretation was not believable, and respondents used terms like “over the top” and “ridiculous” to describe the message. In disbelief, one respondent noted:

“You’re not really going to take that seriously...they’re not going to give this information up if someone came and threatened their life?”

Even when respondents interpreted this statement correctly, they simply did not find it very convincing. We suspect this stems mainly from respondents’ recognition that humans are basically fallible. One person noted that you cannot really tell what particular people will do, even though they may be sworn. Another said the sentence reminded her of the vulnerability of her data, since so many people are involved with collecting and handling it. She summed up her feelings by saying, “To be honest with you, ‘sworn for life’ doesn’t mean anything.” This point was especially germane for respondents who wondered what responsibility former employees had regarding data protection. These respondents recognized that “sworn for life” was a long time, and they were doubtful that former Census employees actually had an obligation to keep

the information confidential once they no longer worked there. One person found it difficult to find any value in this statement from a respondent's perspective; this respondent felt that the sentence was there only because it had to be, that the only people who would pay attention to it would be people who work for the Census Bureau.

To be fair, some respondents were impressed by this language. For example, one respondent exclaimed:

“Wow...they're sworn for life...If anybody had any doubts about who's handling their information, where it's going, how many people are going to be reading this, it tells you they're sworn for life, life can be a long time.”

This type of positive sentiment, however, was heavily outnumbered by negative reactions to this message.

### ***Penalties***

This message appeared as the last statement on the back of Letter A. It mentioned that any census employee who releases your information can be imprisoned up to 5 years, or fined up to \$250,000, or both. Respondents did not always interpret the proper dollar amount when repeating the message aloud; for instance, some recounted the fine incorrectly as “\$25,000” and “\$250.” In addition, respondents varied widely in their reaction to this statement. Respondents had differing opinions about the severity of the punishment. Some thought this was strict and couldn't imagine any circumstances under which an employee would want to release information:

“That you pay a hell of a fine if you screw up and let this information go anywhere.”

In contrast, others did not think the penalties were severe enough, because they believed it would only be a fine, not jail time. In some cases, at least more than the previous two messages, respondents liked seeing this message because it showed that the Census Bureau is serious about protecting respondents' data from abuse.

On the other hand, there were respondents who did not believe that the punishment would actually be enforced. One respondent said:

“If somebody [illegally distributes confidential data], they are probably going to say, ok, it's like in a private high school, they give you the option of you can leave quietly or it goes on your record. Which one would YOU like to do?”

Furthermore, respondents felt this statement was in direct conflict with the confidentiality guarantee on the front of the letter. For example, a respondent stated:

“You're saying you're going to protect us, and at the same time you're saying well maybe there's a leak – we have no control over our employees! So they release it but don't worry, they going to be fined \$250,000.”

This particular respondent pointed out that the penalties statement negated the confidentiality promise because it sounded as if the Census Bureau anticipated employee mistakes, which was evidenced by the fine in place. Some respondents were not reassured by this statement whatsoever because, as they pointed out, once the information is compromised the fines do nothing to protect them. So, in a sense, this “safety check” occurs too late to be of any personal benefit to respondents.

Whether or not respondents seemed to interpret this message correctly, or believe the fines exist to ultimately protect the data and respondents, they did seem to notice this statement fairly consistently on the back of Letter A. We suspect this may be due to the fact that numbers—in this case, “5 years” and “\$250,000”—tend to be a salient feature because it differentiates itself visually from the surrounding text (i.e., numbers versus words), much like a word appearing in bold or italics might attract attention immediately.

### ***“Prevents Anyone from Using Them Against You”***

This message was located on the front of Letter A. Respondents were mixed in their reactions to this statement, although the feedback was positive for most: “Federal law protects the privacy of your answers, and prevents anyone from using them against you.” Respondents generated a number of different scenarios under which data about an individual might be used against them. Some examples included: if you’re housing people and (illegally) charging them, if the number of people living in a housing unit is over the capacity, if you’re a meth lab crack addict, credit card or identity theft, if someone was maintaining two households, or if you’re here illegally without the proper papers. While some of these things could be determined from census data, others seem to be simply things that could cause trouble for people, but are not related to the information contained in the census form.

A minority of respondents had negative reactions. Some respondents simply did not believe the statement:

“It’s like having your information out there in cyber-space – I don’t necessarily believe they will protect it.”

Others felt that the negative tone implied here was scary and they preferred the more positive reference to “rights and benefits” on the back of Letter A.

### ***Privacy Website***

The back of Letter B contained an invitation to visit a website for more information. (“Please visit our website at <http://www.census.privacy.gov> to learn more about our privacy policy and data protection.”) This was consistently well received by respondents. It helped to legitimize the materials respondents read and provided a reference point for those who might have additional questions. Not all respondents said they would go to the website, and many reasoned they didn’t feel the need to search for more information. But, some noted that if they did have questions about the census, confidentiality, or other issues, they thought the website would be a more trustworthy information source than word of mouth.

Respondents noted the following pieces of confidentiality-related information they would like to see included on the website:<sup>12</sup>

- General information about the Census Bureau and what the data are used for;
- Access to census results;
- Explanation of why the information is so important and why confidentiality was stressed;
- Information about privacy policies;
- What technology is used to keep the data safe;
- A link to the law (Title 13) and a description of what it means;
- Administrative record information (what other agencies were contacted, what information is gathered, and the purpose of the data collection);
- Information about how to fill out the form; and
- Who to contact with questions, business hours, and telephone numbers.

### C. Vague Messages

Respondent reactions to these messages are organized under the heading of “vague” because they often had difficulty interpreting these concepts correctly and consistently. In some cases, these terms were of little use to respondents both because they were unable to properly interpret them and they did not seem to have the intended effect.

#### ***“Rights and Benefits”***

The back of Letter A included the following sentence: “The combined data cannot be used to affect your rights and benefits.” Despite the fact that one respondent (quoted in the previous section) preferred this language to the “your answers cannot be used against you” text, the “rights and benefits” statement did not resonate well with respondents overall. Respondents felt these terms were too general or vague. Respondents could think of a few examples of “benefits;” they frequently offered “Social Security” and “Medicaid” when asked for an example. Less frequent mentions included WIC, food stamps, financial aid, and housing, while others couldn’t think of any benefits.

The term “rights” was even less well understood. Many respondents could not give examples (e.g., “just your rights as a person”). Others mentioned “your rights as a U.S. citizen,” “legal rights under the constitution,” and “applying for a job at an agency is your right if you’re qualified.” Many respondents could not think of any examples of rights; this occurred much more frequently than for benefits.

---

<sup>12</sup> Respondents also offered many other topics they would expect to see covered in a Census Bureau website. Please see the recommendation section for a list of topics.

### ***“Oath”***

The reminder postcard contained the following two sentences: “Your information is ONLY seen by people who are sworn for life to protect it. Every employee takes an oath to protect your information.” The first sentence, noted above, was included on the back of Version B without the second sentence.

For the most part, respondents understood the relationship between the concepts of being sworn and taking an oath. Only one person did not understand the word “oath.” However, they interpreted this concept in many different contexts. Some compared it to a notary public, who has the same kind of confidentiality restrictions as implied in the census materials. Others also thought taking an oath was extremely important:

“I assume when these people take an oath to protect your information, they’re doing it as I did it, full-heartedly and to the death, to protect that information.”

But most people thought in less relevant terms, and some were neither impressed nor reassured by this message:

“Judges are sworn to be judges and policemen are sworn to be policemen. So this is really nothing special.”

“Every rent-a-cop takes an oath.”

People did not see much substantive difference between the two sentences. And some thought they were redundant:

“If you’re sworn for life, we assume you’re going to take an oath.”

However, one respondent thought that perhaps not all employees were sworn for life, while the next sentence explicitly states that all employees take an oath. Another thought that, since all employees take an oath, every employee could see your information.

### ***“Dress Rehearsal”***

This concept was present in every version of the letter, but appeared separately from the confidentiality messages; it was embedded in the first paragraph on the front. Even though this research focused on the confidentiality-related material in the letters, respondents often spontaneously commented on the “Dress Rehearsal” concept. They were uncertain what it meant, but many ultimately decided it was a “census test” by reading the surrounding text. It was a phrase that was distracting, and respondents came up with several ways in which it related to the decennial census. One respondent thought it was a sample form, while another thought this was a “pre-screening to make it easier to determine who to send a census form to.” Still another thought they were starting to send the letters out in 2008 to “smooth the road.” One respondent interpreted this concept literally:

“[The letter is] talking about they need help on the census Dress Code, right? They need help on the Dress Code, uhm, the Dress Rehearsal...what is the Dress Rehearsal about, what do they do? Like, what kind of dress? I just wanted to know, is it talking about clothes...?”

#### **D. Letter and Language Preferences**

Of the 36 interviews for which letter preference data were available there was a slight preference for Letter B overall; 12 respondents preferred Letter A, 16 respondents preferred Letter B, and 8 respondents preferred Letter C.<sup>13</sup> Order effects were not apparent when taking into account the sequence of letter exposure. Regardless of the number of people who preferred a particular letter, we observed fairly consistent opinions among respondents regarding particular passages or features across the letters that seemed to resonate with them.<sup>14</sup> It is interesting to note that the interviews conducted in Missouri and Hawaii produced similar findings; we at first thought these interviews would be useful for detecting differences in privacy attitudes and resulting preferences for letters or interpretations of the main messages. By and large, we did not find this to be the case.

Regarding specific wording used in the letters, on the whole, respondents seemed to react positively to the statistical purposes statement and the administrative data uses explanation appearing in Letter B. Respondents tended to prefer the “statistical purposes” statement in Letter B because it seemed very straightforward and clear that the data would only be used to compare “groups” or “averages” for particular variables (i.e., “Your survey answers will only be used to produce statistics, and for no other purpose.”). Surprisingly, respondents adequately and fairly consistently interpreted “statistics” in this context to mean, “to produce numbers...[and] ...groups of information.”

Respondents universally expressed extremely positive reactions to a sentence in Letter A that intended to convey the “statistical purposes” concept, but which was interpreted by respondents as a definition of confidentiality (i.e., “Your answers are confidential. That means the Census Bureau cannot give out information that identifies you or your household.”). An almost identical statement was embedded on the back of Letter C, that is, “Your answers are confidential. That means the Census Bureau cannot give out information that identifies you or your household” and it was interpreted similarly. Respondents mentioned this type of statement was very useful to them because it clarified what the Census Bureau meant when it used the term “confidential” in the previous sentence.

The back of Letter B tended to receive more approval than Letter A when it came to text describing a reiterated confidentiality message and explaining the administrative records use

---

<sup>13</sup> Originally there were 37 usable interviews, however, for one of these interviews letter preference data was not recorded.

<sup>14</sup> These results are fairly consistent with the results from the interviews conducted in Hawaii; they also noted some of the same messages as their favorites, even though the results for letter preference varies slightly from the rest of the interviews. This data was available for 7 of 10 interviews; 3 preferred Letter A, 1 preferred Letter B, and 3 preferred Letter C.

concept. There was evidence that respondents preferred Letter B because it was visually organized in such a way that each topic was contained in its own paragraph (compared to the single paragraph used in Letter A). This approach tends to visually break up the text and perhaps made it easier for respondents to identify and process these various message “packets.” Respondents were dissatisfied with the fact that neither Letter A nor Letter B divulged the identities of the *other government agencies* that might provide additional data about them; however, many respondents reacted positively toward Letter B because they recognized this version at least offered to tell them which agencies would *not* be allowed to access their data. Finally, after seeing the website listed in the last paragraph on the back of Letter B, most respondents remarked they were glad to see a website. Some even said they would expect to find a website on official government communications.<sup>15</sup>

In terms of the overall length of the letters, in relation to the administrative records use message located on the back, respondents preferred longer letters with more detail; 14 out of 22 respondents who received the short versus long letters (i.e., either Letter A or B, compared to Letter C) preferred the “longer” letters. When preferences for the shorter letter (Letter C) were expressed, respondents most often made their decision based on the desire to read less rather than more. Many of these respondents conceded, however, that they did find the text on the back of the longer letters useful.

#### **E. Locating Messages on the Back of the Letters**

Many respondents did not spontaneously turn the letters over to check for or read content that may have appeared on the back in the first round of interviewing. Respondents offered a few explanations for this behavior once they were asked to read the additional text; some indicated they realized there was information on the reverse but were waiting for the interviewer to tell them how to proceed and others said they would not expect any information to be provided there. The latter explanation was supported by respondents who noted that the text on the front of the letter signaled a one-page letter by including the Director’s signature at the bottom of the page.

The finding that respondents might not realize more information was provided on the back of the letter was fairly clear by the end of the first round of cognitive testing. To counteract the possibility that respondents might miss the messages on the letter’s reverse side, a statement was added to the end of the confidentiality paragraph in Letters A and B alerting readers to the presence of additional information: “The back of this letter contains more information about protecting your data.” This statement was pretested in the second round of cognitive interviewing, and respondents consistently navigated their way to the back of the letter and/or reported they were aware of the additional information presented on the back due to this added statement.

---

<sup>15</sup> It is interesting to note, however, that none of our respondents noticed the main Census Bureau website on the front of the letter. We suspect this may have been because they were attending to the text of the letter itself, and not necessarily to the design aspects of the letter. The unnoticed website was located in very small font, on the front of the letter in the lower right corner.



## VI. RECOMMENDATIONS

This section of the report contains a few general observations and recommendations for crafting survey letters, as well as a detailed list of recommendations specific to the initial decennial survey cover letter (See items 1 through 4 on the list below). We also present recommendations for confidentiality- and privacy-related messages across the “mail-out-mail-back” (MOMB) package (i.e., the advance letter, initial survey cover letter, follow-up postcard, and replacement survey cover letter) (See item number 5) as well as some suggestions for future content and development for the Census Bureau’s website (See item number 6).

The results of this research seem to support the findings from previous research undertaken with other survey letters here at the Census Bureau. Successful survey letter messages tend to be those that are as brief, positive, and explicit as possible. In the context of the decennial survey cover letter this means omitting any unnecessary messages that do not advance the main message and perhaps may detract from it. For instance, item number 4b in the list below advises avoiding statements eliciting strong negative reactions (i.e., respondents react in disbelief or feel the messages are “over the top/too far reaching”), as well as avoiding statements that respondents demonstrate inability to consistently interpret correctly (e.g., “72 years” statement). These recommendations are especially critical for messages carried on the front of the letter, a “real estate” location that should be reserved only for streamlined “main” messages. Inserting superfluous statements at this point in the letter tends to distract respondents from the main message.

Additionally, it’s critical to present more detailed information after all of the main messages have been communicated. Certain locations within the letter—primarily the back of the letter—are acceptable locations for such messages. Respondents interpreted information on the back of the letter as the “fine print,” but this perception did not necessarily indicate a negative reaction to the approach. Rather, respondents *expected* the Census Bureau to use this location to include more detailed descriptions of the content located on the front of the letter. Additionally, one design consideration that seemed to counter-balance the “fine print” interpretation was the font size. As noted previously, some respondents spontaneously noticed the text on the back of the letter was the same size as the text on the front, which seemed to dispel any notions that the Census Bureau was trying to hide anything on the reverse of the letter (in contrast to the fine print commonly seen on documents such as credit card solicitations used to communicate the complex terms of the contract). Thus, formatting decisions are critical aspects of the letter’s overall appearance that, like the text itself, convey meaning.

The list of recommendations is as follows:

1. Whichever cover letter version is selected, a field test is highly recommended before including these messages in the 2010 decennial effort.<sup>16</sup> Such a test, with a sample large enough to detect significant differences, could reveal response effects generated by negative (or positive) reactions to these messages.

---

<sup>16</sup> This is also the recommendation of Dr. Eleanor Singer, consultant to this project, and the Statistical Research Division’s staff who carried out the research.

2. The administrative records use message should be located on the back of the survey cover letter. Use the sentence pretested in the second round of cognitive interviewing on the front of the letter to alert respondents to the location of this message (i.e., “The back of this letter contains more information about protecting your data.”).
3. The administrative records use message should not be included on the census form, since there is inadequate space on the back of the form (or anywhere else) to replicate the full explanation recommended for the back of the letter on the form. Based on an earlier recommendation from Dr. Singer, respondent materials should avoid describing the same message in various ways across the materials to keep from confusing respondents about the intended meaning of any particular message.
4. The messages from Letter B are recommended, with the following modifications:
  - a. The definitional statement of confidentiality from Letter A should follow the first sentence (“Your answers are confidential. That means the Census Bureau cannot give out information that identifies you or your household.”)
  - b. Remove statements, concepts, or words that tended to receive “high negative” reactions from respondents, because they were interpreted as potentially frightening, disingenuous (i.e., “over the top”), or were confusing or difficult for respondents to interpret correctly/consistently:
    - “Your census form remains confidential for 72 years.” (on front and back of letters)
    - “immigration agencies” (back of letter)
    - “...only seen by people who are sworn for life to protect it” (on back of letter and reminder post card)

If these recommendations are accepted, the result would be a decennial cover letter that incorporates messages that performed well from the various letter versions we pretested (See below for excerpts of the proposed paragraphs, see Attachment I at the end of this report for the full, recommended letter).

FRONT OF LETTER (confidentiality related messages):

Your answers are confidential. That means the Census Bureau cannot give out information that identifies you or your household. Your survey answers will only be used to produce statistics, and for no other purpose. The back of this letter contains more information about protecting your data.

BACK OF LETTER (administrative records use messages):

**Your Answers Are Confidential:**

Federal law protects your privacy and keeps your answers confidential (Title 13, Sections 9 and 214). This law prevents anyone from using your answers against you. The answers

you give on the census form cannot be obtained by law enforcement or tax collection agencies. Your answers cannot be used in court. They cannot be obtained with a Freedom of Information Act (FOIA) request.

To improve census results, other government agencies may give us additional information about your household. The additional information we receive is legally protected under Title 13, just like your census answers.

Please visit our website at <http://www.census.privacy.gov> to learn more about our privacy policy and data protection.

5. In thinking about implementing pretested confidentiality-related messages (i.e., the confidentiality and administrative records use statements) throughout the MOMB package materials, SRD staff and the working group recommend the following approach:
  - a. The Pre-Notice Letter will not contain confidentiality/privacy related messages (determined to be out-of-scope at the beginning of this project);
  - b. The Initial Cover Letter will contain the pretested confidentiality/privacy messages contained in the attached document (i.e., the third paragraph on the letter's front and all the text on the letter's reverse);
  - c. The Census Form will omit the confidentiality/privacy text on its back page (i.e., third box from the top);
  - d. The Follow-Up Postcard will omit confidentiality messages due to space limitations (the "fold-down" flap design is being replaced by a smaller card with no flap); however, the full suite of confidentiality-related messages are provided in both the initial and replacement questionnaire package; and
  - e. The Replacement Cover Letter should contain the same confidentiality/privacy messages on the front/back--just as in the Initial Cover Letter (to keep consistent messaging across multiple mailings). Other wording may vary (e.g., inclusion of a "mandatory survey" message, further instructions on when and how to return the census form, etc.).
6. An auxiliary recommendation stems from our discussion with respondents regarding their expectations of the website's content—the website appears on the back of the initial survey cover letter. Though the mention of the website merely indicates respondents can find more information about the Census Bureau's privacy policy and data protection, they reported expecting to find a vast array of supplemental information that would help fill in their knowledge gaps where the census was concerned. There is sufficient evidence, however, suggesting that emphasis should remain on concerns respondents voiced more frequently: wanting more information about confidentiality and privacy practices and wanting more information about the administrative records uses activities at the Census Bureau.

We assume these issues could be addressed through the 2010/2008 decennial website (<http://www.census.gov/2010census/>) the Public Information Office (PIO) is currently

developing, either by linking respondents to it from the privacy website or directing respondents to the 2010/2008 website directly; however, the confidentiality/privacy-related content should be placed in a salient location if the 2010 website is given instead of the privacy website.

Below is a list of possible content that could be added to the 2010/2008 decennial website, if it does not already exist. Having reviewed an earlier version of this website, it is clear that much of this content would be new, but much of it exists somewhere on the Bureau's website already. If content is added, care should be taken to use a menu of well-organized links that would carry fairly intuitive names.

- **Additional information regarding confidentiality**
  - What does the Census Bureau mean when it says “your answers are confidential”?
  - What safeguards are in place to keep my data out of the public sphere? (E.g., Title 13/laws, firewall, encryption, employees who have taken an oath to protect the confidentiality of the information, etc.)
- **Information about Title 13 and its significance**
  - What is Title 13? Where do I find it? What does it say (in layman's terms)?
  - How does it protect me (the respondent)? And what makes this particular law so special?
- **Expansion of general information about the decennial**
  - Purpose for the census (e.g., “Uses for Census Data”)
  - Explanation for why the census collects certain types of information (e.g., Why does the Census Bureau collect race?)
  - What is apportionment, and how is the census count used for that?
  - An explanation of the previous uses or benefits from census data being made available
  - Access to previous census data—in a searchable format (e.g., link to “Fast Facts” or “American Fact Finder,” perhaps)
- **Contacts.**
  - A person or office to contact if respondents have questions about filling out their census form or any other aspect of the census (i.e., email, telephone number, mailing address)
- **An explanation of the mandatory nature of the survey**
- **Employment**
  - Information about available jobs/positions
  - Include a list of the Regional and Local Census Offices—RCOs and LCOs

- **An explanation of the significance of the OMB number**
  - It signals a legitimate and sanctioned survey effort
  - When does one use the contact information that is listed with this message, and for what purpose?
- **Provide a contact for language fulfillment**
  - Someone may know an individual who might need materials in a language other than English
- **Provide a facsimile of the census form**
  - This can help legitimize the survey (if people have concerns about whether the census form is a bona fide government document and request)
- **A plain language explanation of the Census Bureau's administrative records programs**

## VII. CONCLUSION

### *Answering the Research Questions*

To the best of our ability, and within the confines of the research project and methodology, we were able to address each of the research questions.

1. *Are respondents able to interpret the main concepts reliably and correctly?* The main message many respondents gleaned from these materials was the idea that the Census Bureau was trying to “do the right thing” by keeping data out of the wrong hands, but they also recognized the Census Bureau could not guarantee a data breach would not occur in the future. Furthermore, some concepts were easier for respondents to interpret correctly and seemed to benefit from the surrounding text. For instance, respondents provided appropriate “guesstimates” of what the Census Bureau meant by “confidential,” and their interpretations were often aided by the sentence immediately following (i.e., “That means the Census Bureau cannot give out information that identifies you or your household”). Respondents, for the most part, also provided consistent (yet vague) definitions for the passage “Your survey answers will only be used to produce statistics...”
2. *How are each of these messages received?* Respondents, as part of the normal course of receiving a participation request, found it fairly customary for a letter such as this to contain a confidentiality pledge. To be believable, however, it is important that the messages avoided sounding disingenuous (or “over the top”) in the process of conveying that information is kept confidential, odd (e.g., “72 years”), or overly bureaucratic and therefore “ignorable” (e.g., presenting the Title 13 reference on the front of the letter). Many respondents claimed to believe the government already shared a great deal of information among various agencies, so they were not necessarily surprised to read the administrative records message. Some noted that this further compromised the confidentiality pledge, but most found this practice to be unproblematic for them personally. Many noted, however, this would not be the case for others who may be in the country illegally or who might otherwise engage in other practices that might cause them trouble should it be discovered (e.g., respondents often mentioned housing a number of people that exceeded the lease agreement or other housing ordinance).
3. *Does any particular message generate negative reactions?* Any messages that elicited puzzled or borderline negative reactions were mainly those deemed “over the top” or those that seemed to further qualify the confidentiality assurance. For instance, in some cases the “72 years” concept (and in most cases the administrative records message) tended to modify the strength of the confidentiality assurance respondents read on the front of the letter. These reactions, however, were not universally represented among all respondents. The practical solution for handling these types of reactions is to exclude those messages, to the extent possible, so as to avoid causing confusion, distraction, or concern for a portion of respondents.

4. *How does the placement of the messages affect respondents' reactions? In particular, how do respondents react to the placement of the administrative records use statement on the back of the letter? On the front of the letter? On the back of the census form? Which placement did they prefer?* Respondents, for the most part, did not object to the administrative records message being located on the back of the letter. A few respondents did mention that "if this information was that important" the Census Bureau should locate it on the front of the census form or on the front of the letter. In most of the interviews, however, the logical flow of information and the order in which it was presented worked well when messages were dispersed across the front and back of the letter. Only about one-third of respondents who read a version of the letter containing all the information on the front preferred seeing this message and no others. Even these respondents noted that they liked having the chance to read more information about the administrative records use on the back of the other (longer) letters. We interpret their sentiments to reflect a tendency to want to read less, rather than more, when they were uninterested in a particular topic. Regarding the placement of this message on the back of the census form, many respondents guessed they might eventually find it, but felt it didn't provide enough information. Regardless of the location of this message, respondents did not perceive the Census Bureau was attempting to hide information. We attribute this finding, at least partially, to the font size used on the back of the letter; it was identical to that on the front of the letter, which helped counteract the interpretation of this message as "the fine print" (e.g., some respondents mentioned credit card statements with truly minuscule print, in contrast to the font used on the back of the letter).
5. *Which letter did respondents ultimately prefer, and why?* Various pieces of the letters resonated with respondents; mainly the statement perceived as the "definition" of confidentiality (i.e., "That means the Census Bureau cannot give out information..."), the revised statistical purposes statement that uses the word "statistics," and the format of the administrative records use statements that named other agencies that would not be allowed to view their data.

The answer to these last two questions was affected, somewhat, by a real change in respondents' perception of risk:

6. *How do respondents feel about the Census Bureau in general?*
7. *Do respondents trust the CB to keep their information confidential and protect their data?*

Their perception was at least in part due to events that were prominent in the media either during or before this research was carried out. The respondents we spoke with expressed high regard for the Census Bureau and many felt that carrying out a census was indeed important work. For these reasons, most respondents indicated they trusted the Census Bureau to collect this information, but they added an important caveat. These respondents almost always concluded by saying "but you can't tell what might happen in the future." This comment was often followed by some discussion of the federal government mishandling and leaking data (i.e., the stolen

Veteran’s Administration laptop) and/or the idea that people in power could change the laws that currently protect their data to allow greater access (e.g., respondents often referred to The Patriot Act, and sometimes the wiretapping stories in the media, to support this claim/concern). From these types of exchanges with respondents we conclude the confidentiality-related messages were not responsible for this pessimistic reaction, rather, these messages seemed to confirm what they already suspected and/or knew to be true. On a positive note, even after recounting several examples of media stories like these, when asked if they would participate in the census, most respondents said they would. So, even though many of the people we spoke with held the Census Bureau in high regard and would likely fill out the form, they were fairly certain the Census Bureau would be powerless to prevent an accidental data leak or prevent a high-ranking political figure from accessing it.

### ***Results in Context—A Range of Application***

Our finding was that respondents, for the most part, were not overly concerned about the government sharing their data. But this finding, regarding respondents’ reactions to the administrative records use messages, may not generalize to other surveys. A similar message, couched in a different survey environment that collected more sensitive or difficult to obtain data, could cause respondents to react more defensively about providing their personal data in response to a survey or census.

We suspect that one of our results may have been highly influenced by an exercise we embedded in the interview protocol, which may have created a more proportionate response to the significance of the data collected by the decennial form and the idea that government agencies share their personal information. Prior to allowing respondents to read the letters, we showed them the census form and briefly reviewed each type of data it requested, as we explained the letters would be mailed along with the form.

The lack of an extremely negative reaction to this message was partly due to the fact that many respondents already thought the government engages in this type of activity. Some seemed to find this concept relatively unproblematic because they had seen the census form and felt the information it collected was not very “private,” and that this type of basic demographic information was already widely available to the public (e.g., “This is all [the census form] asks? You could practically get all this information off of the internet.”). In fact, many respondents spontaneously offered they were surprised the information requested was not terribly detailed. These may be reasons why extreme negative reactions were kept at bay. This finding runs contrary to earlier work carried out with survey letters in the SIPP environment; respondents were far more concerned about data sharing after reading a similar message, perhaps because they were not given the opportunity to see the survey questions (Landreth, 2002). Thus, care should be taken to research the reactions of respondents viewing this message within the context of more sensitive data collections.

### ***Future Research***

Additional research with these types of messages could help clarify some issues that seem confounded in this project. First, it would be useful to be able to disentangle the effects of “boilerplate” messages—those being jargon-laden legal statements appearing ubiquitously on all government documents—from the true saliency of the administrative records use message. That



is, we wondered whether the lack of reaction on the part of respondents was because they ignored, or glossed over, these types of passages or because they did not necessarily process the message very deeply (it was surrounded by a fair number of additional messages). More work in this area might involve using eye-tracking equipment to detect the amount of time respondents spent reading specific messages. Coupled with cognitive probes, this research would shed light on how well respondents process and perceive messages regarding the Census Bureau's statistical activities.

Second, it would be worthwhile to study the effect of the methodology on the saliency of these particular confidentiality-related messages. We strongly suspect that the introduction of the census form affected respondents' interpretation of these messages. More research could be done to compare the effect of the methodology under conditions we implemented versus allowing respondents to review the confidentiality-related messages under a condition that allows them to navigate the mail-out-mail-back (MOMB) package without the direction of the cognitive interviewer.

Third, future cognitive research of this type should account for respondents' reading level capabilities and its effect on comprehension. For this study we used education attainment level as a loose proxy for reading level, but we had no way to judge actual reading level. This method is far less precise than we would like, and it would be useful to develop a measure for future studies would help determine which concepts cause the most comprehension difficulties at lower reading levels and how interpretation is affected.

Finally, a field test would be required to fully understand the influence certain design choices, regarding the presentation of the administrative records use messages, have over response rates. The inclusion/exclusion of the administrative records use paragraph should be experimentally manipulated, as well as including this particular message on the front versus the back of the letter.

## EPILOGUE

Since the original research was conducted and presented to the Working Group and sponsor (the Decennial Management Division, or DMD), several changes to the letter have developed.

First, upon the original presentation of the results from this research, the sponsor accepted the recommendations and began routing the recommended letter through the correspondence approval process. Minor changes (edits) were made to the letter that did not affect the privacy- and confidentiality-related messages.

Second, in consultation with the members of the Working Group, it was decided the “72 years” concept needed to be included because it explains a limitation to the confidentiality statement/assurance. In light of the findings for this message, the Working Group and the internal and external consultants decided an alternative message (that explained the data were released eventually and offered an example of the data uses at the time of release) was appropriate for the back of the letter, where it would be less likely to distract respondents from the main confidentiality message. See Attachment II for the exact text.

A version of the letter incorporating these first two developments is reflected in the Attachment II (dated 3/15/2007).

Finally, after reviewing the proposed letter’s contents, executives on the Data Stewardship Executive Policy Committee (DSEP) concluded the description of the administrative records use program was incomplete and should not be included in the 2008 Census Dress Rehearsal mail materials. In a meeting in May of 2007, the DSEP elected to remove this paragraph until a future research effort could be mounted to explore the appropriate contents for such a message.

## REFERENCES

- Fay, R., N. Bates, and J. Moore (1991) "Lower Mail Response in the 1990 Census: A Preliminary Interpretation." In Proceedings of the Annual Research Conference. Washington, D.C.: Bureau of the Census.
- Frey, J. (1986) An Experiment with a Confidentiality Reminder in a Telephone Survey. *Public Opinion Quarterly* 50:276-9.
- Gerber, E. (2002) Privacy Schemas and Data Collection: An Ethnographic Account. Census 2000 Testing and Experimentation Program Report.
- Gerber, E. (2003) "Respondents' Understanding of Confidentiality Language," Proceedings of the American Statistical Association, Survey Research Methods Section [CD-ROM], Alexandria, VA: American Statistical Association.
- Guarino, J., J. Hill, and H. Woltman (2001) "The Effect of SSN Requests and Notification of Administrative Record Use on Response Behavior in Census 2000," Proceedings of the American Statistical Association, Survey Research Methods Section [CD-ROM], Alexandria, VA: American Statistical Association.
- Kerwin, J. (1998) Cognitive Testing of Privacy Statements in Census 2000 Dress Rehearsal Letters. Internal report completed under contract by Westat. Washington, DC: U.S. Census Bureau.
- Landreth, A. (2001) SIPP Advance Letter Research: Cognitive Interview Results, Implications & Letter Recommendations. Center for Survey Methods Research, Statistical Research Division's Study Series, Survey Methodology #2001-01. Washington, DC: U.S. Census Bureau. Last accessed on May 3, 2007 [<http://www.census.gov/srd/papers/pdf/sm2001-01.pdf>].
- Landreth, A. (2002) Report of Cognitive Interview Results from the Alternative Administrative Records Use Paragraph of the SIPP Advance Letter. Center for Survey Methods Research, Statistical Research Division's Study Series, Survey Methodology #2003-11. Washington, DC: U.S. Census Bureau. Last accessed on May 8, 2007 [<http://www.census.gov/srd/papers/pdf/ssm2003-11.pdf>].
- Martin, E. (2001) Privacy concerns and the census long form: Some evidence from Census 2000. 2001 Proceedings of the American Statistical Association, Statistical Computing Section [CD-ROM], Alexandria, VA: American Statistical Association.
- National Research Council (1979) Privacy and Confidentiality as Factors in Survey Response. Washington, DC: National Academy of Sciences.
- Singer, E. (1978) "Informed Consent: Consequences for Response Rate and Response Quality." *American Sociological Review* 43(2): 144-162.

Singer, E. "Public Perceptions of Confidentiality and Attitudes Toward Data Sharing by Federal Agencies." 2001. Pages 341-370 in Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies. (Editors: Pat Doyle, Julia Lane, Jules Theeuwes, and Laura Zayatz). New York: Elsevier.

Singer, E., H. Hippler, and N. Schwarz (1992) "Confidentiality Assurances in Surveys: Reassurance or Threat?" International Journal of Public Opinion Research 4(3): 256-268.

Singer, E., J. Van Hoewyk, and R. Neugebauer (2003) Attitudes and Behavior: The Impact of Privacy and Confidentiality Concerns on Participation in the 2000 Census. Public Opinion Quarterly 67: 386-384.

ATTACHMENT I  
*(Recommended Survey Cover Letter—front of letter)*

DATE XX, 2007

Dear Resident:

The United States Constitution requires a census of the United States every 10 years. To prepare for the 2010 Census, the U.S. Census Bureau is conducting the 2008 Census Dress Rehearsal. The results will be used to develop better methods that will make the 2010 Census easier, more convenient, and less costly for taxpayers.

**Please complete and mail back the enclosed census form today.** We need your help to improve the 2010 Census. Results from the 2010 Census will be used to help each community get its fair share of government funding.

Your answers are confidential. That means the Census Bureau cannot give out information that identifies you or your household. Your survey answers will only be used to produce statistics, and for no other purpose. The back of this letter contains more information about protecting your data.

Sincerely,

Charles Louis Kincannon  
Director, U.S. Census Bureau

Enclosures

*(Recommended Survey Cover Letter—back of letter)*

### **Your Answers Are Confidential**

Federal law protects your privacy and keeps your answers confidential (Title 13, Sections 9 and 214). The answers you give on the census form cannot be obtained by law enforcement or tax collection agencies. Your answers cannot be used in court. They cannot be obtained with a Freedom of Information Act (FOIA) request.

To improve census results, other government agencies may give us additional information about your household. The additional information we receive is legally protected under Title 13, just like your survey answers.

As allowed by law, census forms become public after 72 years (Title 44, Section 2108). These forms can be used for family history and other types of historical research.

Please visit our website at <http://www.census.gov/privacy/> to learn more about our privacy policy and data protection.

ATTACHEMENT II  
*(March 15, 2007 Version of the Survey Cover Letter—front of letter)*

DATE XX, 2007

Dear Resident:

The U. S. Constitution requires that a census be conducted every 10 years. To prepare for the 2010 Census, the U.S. Census Bureau is conducting the 2007 National Census Test. The results will be used to develop better methods that will make the 2010 Census easier, more convenient, and less costly for taxpayers.

**Please complete and mail back the enclosed census form today.** We need your help to improve the 2010 Census. Results from the 2010 Census will be used to help each community get its fair share of government funding.

Your answers are confidential. That means the Census Bureau cannot give out information that identifies you or your household. Your survey answers will only be used to produce statistics, and for no other purpose. The back of this letter contains more information about protecting your data.

Sincerely,

Charles Louis Kincannon  
Director, U.S. Census Bureau

Enclosures

*(March 15, 2007 Version of the Survey Cover Letter—back of letter)*

**Your Answers Are Confidential**

Federal law protects your privacy and keeps your answers confidential (Title 13, U.S. Code, Sections 9 and 214).

The answers you give on the census form cannot be obtained by law enforcement or tax collection agencies. Your answers cannot be used in court. They cannot be obtained with a Freedom of Information Act (FOIA) request.

To improve census results, other government agencies may give us additional information about your household. The additional information we receive is legally protected under Title 13, just like your survey answers.

As allowed by law, census forms become public after 72 years (Title 44, U.S. Code, Section 2108). These forms can be used for family history and other types of historical research.

Please visit our website at <http://www.census.gov/privacy/> to learn more about our privacy policy and data protection