



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

AUG 30 2016

MEMORANDUM FOR: Barbara A. Anderson
Chair
Census Scientific Advisory Committee

From: John H. Thompson
Director

Subject: Census Responses to the Census Scientific Advisory Committee
(CSAC) Recommendations

The U.S. Census Bureau thanks the CSAC for the recommendations submitted as a result of the April 14-15, 2016, meeting.

Your expertise is necessary to ensure that the U.S. Census Bureau continues to provide relevant and timely statistics used by federal, state, and local governments as well as business and industry, in an increasingly technologically oriented society.

Attachment:
Census Responses to the CSAC Recommendations from Spring 2016 Meeting

U.S. Census Bureau Responses to the CSAC-Census Scientific Advisory Committee 2016 Spring Recommendations

1. PRIMUS

1.a. There were several concerns about PRIMUS. Some consist of questions and some of comments and recommendations. It is not clear whether CSAC input to PRIMUS could best occur through the formation of a Working Group, or just to have some CSAC members with particular expertise talk with the PRIMUS staff. However, the section below discusses CSAC's concerns and questions. PRIMUS will be the way that *most* Americans interact with the 2020 Census. We hope that the standard for evaluating the success of PRIMUS will be not only a high response rate, but also accurate responses and a positive user experience. The 2016 PRIMUS test appears to have not given enough consideration to user experience. While the Census Bureau will inevitably need to do some post-processing of the data, such as correcting errors and removing duplicates, careful design will ultimately reduce costs and errors, increasing data quality and user experience. We would like to know about any plans to incorporate user-centered design methods into PRIMUS development. We also would appreciate a summary of the user experience and survey design research conducted on electronic versions of the decennial form, for PRIMUS, but also prior to PRIMUS (since 2010). We would like a response from Census Bureau about the planned process for decision making for the 2017 PRIMUS test. What is the timeline? What is the planned evaluation of the 2016 test? How exactly is feedback being gathered from survey design and user experience experts? Who is making decisions about potential changes in response to that feedback?

CSAC requests a webinar this spring to learn more about PRIMUS.

Census Response:

On May 26, 2016, the Census Bureau announced a design decision that has been made in support of the Census Enterprise Data Collection and Processing Program. Since December 2014, the Census Bureau has been assessing whether to use commercial software products to collect and process data, or whether to continue building our own systems. After a great deal of evaluation and discussion, the Census Bureau has determined that a hybrid approach – combining a commercial off-the-shelf (COTS) system with specific solutions developed by Census Bureau experts – will best meet our needs.

The COTS platform we have selected supports capabilities proven to be viable in the marketplace for many customers including other Federal agencies. With this decision, the Census Bureau will integrate the platform into the Decennial Census operations, beginning with the 2017 Census Test, and then into all census and survey operations. Accordingly, the COTS platform will take the place of the in-house Internet data collection instrument (Primus), and the Census Bureau will continue its efforts in continued innovations and helping find new efficiencies as we integrate this platform into our operations.

We are actively working on increasing input, feedback, and participation in the development of the Internet self-response application as we move towards the 2020 Census. Many of the recommendations made by CSAC members on the Internet self-response application are currently being researched, as we begin configuring the instrument for the new platform. We also plan continued rigorous cognitive and usability testing for the 2017 application.

1.b. Although we hope there is a more systematic assessment of Primus from a user experience perspective, some immediate design issues jumped out.

Census Response:

Due to the Census Bureau’s decision on a hybrid approach, the responses to the Primus related questions are provided in **Table I** (for questionnaire content and design issues) and **Table II** (covering potential glitches).

Table I. Primus		
Number	Comment	Census Bureau Response
1 (Content & Design)	The content needs to be adapted for the online implementation. Basic example: no need to shorten “African American” to “African Am.” Since there’s not the same space constraint as on the paper form. More generally, there are many places where the survey does not adhere to web survey design best practices—for example, the font is too small, it uses bolding rather than underlining and color to highlight text, screen layout is not maximized for space, radio buttons/clickboxes are too small to navigate on a mobile device (it’s better to make the label clickable and use buttons instead). Overall, many of the instructions were wordy with undifferentiated text and limited visuals.	<p><i>The 2016 Census Test will provide the opportunity to test a combined race/ethnicity question with variations in wording for the Black or African American racial category to compare the use of “American” with the abbreviated “Am.” (“Black or African American” v. “Black or African Am.”). It will also explore whether or not the wording change has adverse effects on the ways the Black or African American population responds to the race/ethnicity question. In addition, the research aims to provide insights to how respondents identify with the race category, depending on the wording that is used to describe the category itself (“Black or African Am.” v. “Black or African American”). The results of this research will inform how the category will be displayed for the 2020 Census. We would be pleased to update the Committee on the results of this research.</i></p> <p><i>Our testing for 2017 will include enhancing the overall layout and navigation of the instrument and user experience.</i></p>
2 (Design)	It would be better to have immediate feedback for a field (e.g., if something is entered in a format that is not acceptable, let the user know right away). Right now, it is not always clear which field is creating the problem or if it is a particular subset of conflicting answers across fields. Only sometimes did the error prompt flag the specific issue.	<i>We agree, and in 2017, we plan to test the edit messages, especially those most likely triggered based on the paradata from past Census Tests.</i>
3 (Design)	It is frustrating to have a separate text box for street number and street name. How many people type in both only to have to edit once they notice the next field asks for Street Address? If that’s not feasible, it would still be helpful to shorten the text box and include “#” in the field to help make clear the information that is expected. Roberto suggested that the full address be a single field (e.g., with Google).	<i>With respect to recommendations on the address fields, it is necessary to keep the house number, street name and apartment number as separate fields, in order to better ensure accurate address matching in Census Bureau systems. However, the data entry field will be shortened for both the address number and the apartment number to illustrate the distinction from the street name field.</i>

4 (Content)	<p>Relatedly, instructions and examples that are used on the paper form are not necessarily ideal for online implementation. For example, on the race question, including specific examples for each category clutters the screen and means that not all response categories are visible on one screen (at least when taking on my iPhone)—a survey design no-no. In this case, you might consider providing a link to examples rather than including them next to the categories. Other times, instructions could benefit from elaboration. E.g., perhaps “home equity line” could use further context or elaboration. It is also possible to give clearer indication of what information is being requested (e.g., “insert number” or “#” could be inside text box for number of members of household) and the text box size is currently too large for an entry that is no more than a double digit. Remember that many people will not read the instructions. Make it easy for them to give the right answer without detailed reading of instruction, but also make it possible for them to find the answer to their questions if they have one (e.g., why is a question asked).</p>	<p><i>The Census Bureau research over the past several decades has shown that the use of examples in the race and Hispanic origin questions aids respondents in reporting detailed Hispanic, Asian, and Pacific Islander groups (Cresce and Ramirez 2003; de la Puente and McKay 1995; Gerber et al. 2005; Martin et al. 2004; Martin 2006; McKenney et al. 1988). Our goals for testing examples are to clarify the meaning of checkbox response categories for respondents. Based on positive results with the 2010 Alternative Questionnaire Experiment, our research for the 2015 National Content Test (NCT) examined ways to implement examples. However, we do not want respondents to think that they are limited in choosing only among the examples provided, and that the list is exhaustive. Our research led us to select six detailed example groups for each of the major race/ethnic categories.</i></p> <p><i>We agree that providing a link to a more complete/comprehensive list of examples is an interesting idea. In fact, the 2015 NCT did include reference materials for respondents to utilize when answering the web-based versions of the questions. This included “Help Text” to assist respondents with understanding the definitions of the race/ethnic categories, and an explanation of the question and their reporting options.</i></p>
5 (Design)	<p>It could be helpful to include the Census phone number in the footer of every screen. Allow someone to complete the census form on the phone if they call. The Census logo does not currently link to the Census Bureau website – might be nice to include in case someone wants to learn more.</p>	<p><i>In 2017, we will be researching the feasibility of including a link to the Census Bureau website on the introduction screen. Including a link on each page of the survey would need to be tested to determine whether this would lead to increased numbers of people who leave the survey via the census link and never return to complete it. In addition, placing the phone number to each screen, while being user friendly, would need to be tested to determine how much this would decrease Internet response and increase the cost for Census Questionnaire Assistance.</i></p>
6 (Design)	<p>On the paper form, respondents can see how many questions they need to answer. On the web, it is not clear, which can be frustrating for the respondent. You could say that you need to answer 10 questions about each person in the household. Or you could give the 10 minute time estimate on the screen rather than in a FAQ.</p>	<p><i>We agree, and in 2017, we plan to add clearer guidance on what questions and topics to expect in the survey.</i></p>

7 (Content)	<p>The race questions, in particular, do not translate well to the online mode. It's not clear why only a few select categories are offered. I'm concerned that the question could offend given the large number of hyphenated Americans, the potential that some won't know country of origin, and the sense that some categories are being prioritized (there is certainly room to add more categories so it appears it is a deliberate choice). I thought the Census Bureau had done some more sophisticated testing of this question, but that does not appear to be reflected here.</p>	<p><i>Please see the related discussion and response to Recommendation #4 (above) regarding the selection of categories and the use of examples. The 2015 National Content Test is our primary vehicle for testing the race and ethnicity questions for the 2020 Census. The 2015 NCT is being analyzed this year and the results of this research will inform recommendations and decisions about the 2020 Census content.</i></p>
8 (Content)	<p>Reconsider the order of the questions so that easier questions are asked first. So, ask sex and age before asking about relationship status. Since it's a really long list, is it possible to use age to reorder according to most likely categories? It would be nice to have tabs in which respondents could toggle between the different questions.</p>	<p><i>We have historically asked relationship first to establish the head of the household, followed by sex and age (along with date of birth) in the Decennial Census and the American Community Survey (ACS). Changing this question order is not something we are testing right now, but we understand the purpose of this recommendation.</i></p>
9 (Content)	<p>Reconsider the basic structure of the survey that separates the household and individuals (with individuals completed separately) – it could be more user-friendly to ask individual items and only then ask if the house is owned or rented. Or you might want to answer the age of all individuals in the house at one time.</p>	<p><i>The tenure question is used to determine the reference person used in the relationship question. We are not currently planning to relocate the tenure item to the end of the interview. We considered this before the 2010 Census and decided that if we moved it to the end of the questionnaire, after all the person items were answered, it might be missed (particularly if they break off in the middle of the person questions). We still believe this to be the case, but appreciate the recommendation.</i></p>
10 (Design)	<p>Instead of including PO Box as an option on the first page, instead you could have a "Do not have a street address" that branches to options such as PO Box, homeless, temporary housing, etc. Seems like current structure could encourage people to pick PO Box even if they have a street address and it's not clear what to do if you are homeless.</p>	<p><i>Regarding the recommendations on the address options, in 2017 and beyond, there will no longer be radio buttons to toggle between the three address types. Instead, the default will be to permit entry of a city-style address (address number, street name, etc.) and the option to provide "other address" if the default address fields do not apply.</i></p>
11 (Content)	<p>It could be frustrating to have two questions asking if there are additional individuals in the household. It was easy to miss that one question asked about living there and the other asked about people staying there. Again, it could be more user-friendly to gather information about individuals listed and only then ask about others in the household.</p>	<p><i>The 2016 Census Test tests different approaches to asking about additional people who might have been left off the initial household roster. Some respondents saw a two-question approach (which is referred to in the comment above) that included one question that listed multiple examples of types of people that are often missed (undercounted), followed by a separate question that asks specifically about people with no permanent place to live. Other respondents saw a one-question approach that combined the two questions by including people with no permanent place to live among the larger list of people who might have been missed. For the 2017 Census Test, we will move forward with the one-question approach, taking into account reducing respondent burden, getting the household count correct, and minimizing the number of screens for electronic modes.</i></p>

12 (Design)	On the map, it's confusing to have an outline with dots—it appears as if it is a request to pick a specific house address rather than a block. On an iPhone, the map was especially difficult to navigate—it wasn't clear what was expected because a default shaded area did not show up.	<i>Regarding the map interface, the Non-ID Processing team will be making a preliminary recommendation regarding its inclusion in the 2020 Census design, following analysis of results from the 2016 Census Test and compared to results from the 2015 Census Test.</i>
13 (Design)	One idea for reducing duplicates is to ask at the end of the survey if the respondent would like to notify a member of their household that the Census form had been submitted. There could also be the option to post on Facebook or Twitter. Of course, keep in mind that you want it to be posted with a positive user experience. Otherwise it could actually flame fires of discontent (e.g., about race question).	<i>We will consider this idea as we plan our 2017 testing.</i>
14 (Design)	One idea for improving the respondent experience is to include some campaign messaging in the survey (motivation for participating)—for example, the first screen could give a sentence about the importance of the survey and that completing it online saves government money; when the survey is completed, there could be a thank you message that reinforces this. In terms of messaging around the content, do the FAQ actually cover the questions people have about the survey and about completing the survey online? It is lengthy—it's unfortunate that the "close" button requires scrolling to the bottom (although we eventually realized you could also click anywhere to close. Perhaps add an "X" in the upper right corner?)	<i>We appreciate these recommendations and will consider them as we plan enhancements to the instrument for 2017.</i>
15 (Design)	Census should be measuring respondent experience (satisfaction, burden, confusion) for at least some subset of respondents.	<i>We will look for opportunities to do this as we develop our 2017 testing plans.</i>

Table 2. Potential Technical Glitches	Census Bureau Response
<ol style="list-style-type: none"> 1) Answering on an iPhone on cellular data, the website froze. 2) When one address was put in and then replaced with a different address, the original address was the one that was saved. 3) Taking on the iPhone, the user is several times warned that there were 2 minutes left before timing out. It wasn't clear that the time was reset by clicking OK. 4) Sometimes an error prompt gave information about the problem, but other times it did not. The system allows people to give an age that doesn't match to date of birth. 5) In the demonstration, the PO Box option appeared not to function correctly (it was after we had tried another address, so perhaps it was an issue with how data were being stored/discarded). 	<i>The Census Bureau will address potential technical glitches noted in the recommendations</i>

1.c. Request: It would be very useful to have more devices/computers available during the next meeting and/or provide other ways of accessing prototypes so more of the committee can interact with them and provide feedback.

Census Response:

The Census Bureau received the Committee's request for additional devices/computers during the next meeting and we will evaluate this capability for the next meeting.

2. 2020 Decennial Census Plans

2.a. 2020 Census Program Overview: CSAC was interested to hear about the instability problem with cloud storage and thinks that solution to this problem should be a top priority. CSAC would like to hear an update on this at the next meeting.

Census Response:

We appreciate the Committee's interest and we will update the Committee on our progress at the next meeting.

2.b. Census Test in Puerto Rico: The census returns from Puerto Rico may be atypical at this time, because of the extreme economic conditions on the island and the depopulation that is occurring (particularly in younger age groups). For example, there may be a greater than normal tendency for parents to report children to be resident with them, when they are on the mainland seeking employment, because parents are hoping/expecting their children to return to Puerto Rico when the economy improves. In particular, complications may arise in Puerto Rico from the extreme economic downturn that are potentially different quantitatively and qualitatively from having to correct for the normal migration patterns on the mainland.

Census Response:

The Census Bureau appreciated the comments on Puerto Rico and will consider the circumstances as we prepare for the 2017 Puerto Rico Census Test.

2.c. 2020 Census Partnership Program Question: What changes (if any) are envisioned to the partnership program given the internet option? Also, it would be good to know about Census plans for promotion via social media. Will people be able to go directly to a web site to fill out their survey? It would be good to have a presentation about the partnership program and social media plans for Census 2020 at the fall or spring meeting.

Census Response:

The Partnership Program will use many activities and initiatives from the 2010 Partnership Program. Many of these activities are not mode dependent. However, the messages partnership staff will use when speaking and in the educational and promotional materials will focus on the importance and benefits of participating in the Census, as well as the easy options by which to respond, with a heavy emphasis on online response.

The Partnership Program was implemented as part of the 2015 and 2016 Census Tests and will be a part of the upcoming 2017 Census Tests and 2018 End-to-End Census Test. During the 2015 and 2016 Census Tests, partnership staff attended many community events and activities. The staff had tablets they used at these events for participants to go online and respond. People

reacted very positively to this. Additionally, as we plan for the 2017 Puerto Rico Census Test, we are planning to implement a trusted voices mobile outreach van initiative. These vans will be equipped with tablets to allow people to respond. These are a couple of examples of how some changes will be made with the consideration of the internet option.

The Integrated Partnership and Communications (IPC) Program will use social media as one of the main mechanisms to communicate with the public for the 2020 Census. Once the integrated communications contract is awarded in August 2016, we will work with the contractor to develop plans on how we will implement social media, as well as many of the other IPC components. Currently, we anticipate completing this plan in Spring 2017. The Census Bureau is currently developing the Partnership Program plans. The Census Bureau will present this plan, which includes social media to CSAC.

2.d. Results of the 2016 and 2017 Census Tests: It is suggested the Census Bureau release the 2016 and 2017 data for limited or complete user community to examine and give feedback to the Census Bureau. This is especially important since the implementation of the 2020 Census will be substantially different from earlier censuses.

Census Response:

The Census Bureau appreciates your suggestion to release the 2016 and 2017 data for research. We will investigate the feasibility of providing access to our test data through the Census Bureau Research Data Centers.

2.e. Back-Up Plans: It would be good to know what back-up plans the census is putting together in case some parts of the implementation plan do not work out as planned. This would likely be part of the risk assessment.

Census Response:

The 2020 Census Program Risk Review Board has worked with the decennial subject matter experts to develop risks associated with the 2020 Census program risks. The 2020 Census Risk Review Board also has identified mitigation and contingency planning needs based on those risks. We would be happy to brief the committee in more detail at a future meeting.

2.f. Non-ID Processing and Duplicate Answers: CSAC would like to know more about the implementation of the non-ID processing from the 2016 test, including how duplicate answers will be identified.

Census Response:

The Census Bureau utilizes what is known as the Primary Selection Algorithm for these purposes. Releasing details of exactly how this algorithm works could open the door for fraud, so it is considered a statistically sensitive component of our data processing, with controlled exposure only to those with a programmatic need to know. However, once we complete our analysis of this operation for the 2016 Census Test, we will share the findings with you.

2.g. 2020 Census: Reengineering Address Canvassing

CSAC recommends that Census extend agency expertise in automated image processing, change detection, use of ancillary data, and feature extraction to leverage and enhance automated delineation of residential areas and subsequent extraction of, and reasoning about, changes to residential (address)

structures. The extension could take the form of full-time staff, post-doctoral researchers, visiting scholars or long term partnerships with the private sector or academic institutions.

The Census should seek assistance in integrating automated processes with the currently developed (mostly manual) workflow. The objective is to expose system developers at Census to the potential benefits of full automation, to develop methods that can automate reliably, and to implement capabilities either in-house or through outsourcing that can significantly reduce the need for in-field address canvassing, beyond the current estimates of a 10% reduction. Oakridge Labs and Zillow/MLS could be source of potential help. The REIS data set on commercial-to-residential conversions could also be helpful.

Comment: Given the large and high quality work being done on address data management, the Bureau should seek to figure out how to decouple the address file with (x,y) coordinates from the title 13 restrictions. The FGDC is currently assessing who will lead addressing for the nation and Census is the logical one to do it. There are very large financial benefits and good government results that could result from Census taking leadership and owning "A 16" leadership here.

Census Response:

The Census Bureau agrees that an automated change detection process will provide benefits to the In-Office Address Canvassing methodology. Geography Division staff are developing a prototype automated change detection process as a means to identifying the ways in which the process might integrate with and support the interactive In-Office Address Canvassing methodology currently in operation as well as to identify requirements for potential collaborative work with external researchers and organizations. Geography Division's research and prototyping has focused primarily on automated change detection using imagery and other remotely sensed data, such as the National Land Cover Database, but also has explored potential use of Light Detection and Ranging (LiDAR) data. Our plans call for reaching out to experts in industry and academia as well as Oak Ridge National Laboratory to help chart an appropriate course for integrating automated change detection into the In-Office Address Canvassing process.

*The Census Bureau appreciates the comment related to decoupling the address file from Title 13 restrictions. In the *Baldrige v. Shapiro* ruling (February 1982), the Supreme Court majority held that addresses for individual housing units were information collected about individuals as part of a census or survey. Therefore, the Court ruled that addresses and address lists collected and used by the Census Bureau were specifically protected by Title 13. Absent new legislation that alters the language of Title 13 to exclude addresses from the restrictions on data confidentiality, the Census Bureau cannot share address or lat/long coordinates.*

At the May 31, 2016 meeting of the Federal Geographic Data Committee Executive Committee, it was suggested that the Census Bureau and the Department of Transportation would be logical co-leads. Discussions are underway on this matter and we would be pleased to update the Committee as this develops.

2.h. 2020 Census Program: Optimizing Self-Response

The cost to the Census of having to implement as many as 5 or 6 response attempts to contact a non-responder is high. Perhaps there is some financial or non-financial incentive that could be offered to encourage more non-responders to respond, such as offering free coffee at a local Starbucks or taqueria (where Census takers would be available) or some free US stamps. Of course, there is the possibility that

those who would otherwise have responded will now choose to not respond, in order to receive the incentive.

2.i. One could include a box on the digital form that would permit people, if they wish, to receive via email a link to the Census info on their Census tract/city/state when the Census is complete. People might feel better about filling out the form if there were some personal value-added they received once the Census was complete.

Census Response:

Exploring the use of incentives to increase participation has not been studied during our early research and testing efforts for the 2020 Census. Although the Census Bureau does use incentives to try and increase response rates for some of its lengthier and voluntary surveys, we have never done this for surveys or censuses for which response is mandatory by law, as is the case for the Decennial Census. The Census Bureau and Department of Commerce considered the incentives idea prior to the 1990, 2000, and 2010 Censuses, and each time decided not to pursue it. A key factor in the decision each decade was that response to the Decennial Census is mandatory by law, and that we should not (and probably cannot legally) offer payment to do something that is required by law. A number of other similar operational and financial issues surfaced in these considerations each decade as well. For example, how would the payment of the incentives be operationalized—e.g., As something included with the mailing package being delivered to every address in the nation? Upon proof of response—who would manage that to ensure it was done correctly? Since there would be costs of the incentives and of developing/managing the effort, another issue raised each decade was whether response rates would increase enough to offset those costs.

With respect to including an option in the Internet self-response application for respondents to request confirmation, the Census Bureau is currently exploring this added functionality for future iterations of the application.

2.j. 2020 Census Program: Utilizing Administrative Records

It is suggested the Census Bureau consider using spatially aggregated demographic segmentation data (commercial data such as Tapestry or Claritas data) aggregated to census block groups or zip codes as additional administrative data. This would be a way to further characterize general behaviors of neighborhoods and households. This would be useful in understanding geographic behaviors such as response rates and understanding the best methods for reaching these populations.

It would be good if the Census Bureau managed to get utility companies to share their data. A Census enumerator is seldom the first official to visit a dwelling. Perhaps the Census Bureau could find some utility companies to pilot a data-sharing arrangement with. If appropriate, Peter Glynn might be able to make inquiries in Vermont, and Kathy Pettit could make inquiries with local data intermediaries who regularly access utility data.

It would be good to see more detail on the algorithms that are used to determine when administrative records are used to eliminate NRFU. The Census Bureau indicated that some information was available in the November 2015 webinar. A session on this at the next meeting would be helpful.

CSAC would like to know when the Census Bureau would finalize the list of administrative records that can be used. It would be wise to set a deadline by which a final decision is made as to what records will be used in 2020.

If an individual respondent can be uniquely identified and matched to a previously collected record, then the census has valuable data to assist in the update and/or to use for real-time validation/QC. The update could be more of a database transaction versus a complete rebuild of a new database. This is an area where the census should spend time planning for innovative changes to the process of collecting data. This is a major change in the process, and it will have many challenges (including data protection related to title 13). As such, this is a visionary process that could potentially be integrated in census processes beyond 2020.

If an individual could be uniquely identified and a secure process in the use of administrative data could be implemented, then the census data collection could resemble a process like Turbo Tax. In other words, many questions could be pre-populated and simply ask for validation. If a secure process cannot be established to prepopulate questions for validation, use of data matching real time would still be beneficial. In this case, the respondent would not even have to know that the record is being matched and checked in real time. Data entry could be checked for address spelling errors, date entry errors, etc. Areas where there were discrepancies could be flagged for resolution. This process would add to the quality of the data collected. The Census has data from previously collected surveys to prepopulate a substantial portion of a new Decennial Census, provided some sort of validation process like matching individuals for validation.

Census Response:

The Census Bureau appreciates your thoughtful response and recommendations regarding the utilization of administrative records in the 2020 Census. As documented in the 2020 Census Operational Plan, the Census Bureau has reached decisions on the core set of administrative records and third-party data it will utilize to reduce the Nonresponse Followup workload. The final list of administrative records and third-party data that will be used in the 2020 Census will be announced by September 30, 2018.

We agree with your recommendation on providing additional detail on the algorithms we are using to reduce the 2020 Census NRFU workload and look forward to providing this information in a session at the next meeting.

As we are thinking about the 2030 Census, your recommendation to consider approaches and opportunities for an individual respondent being uniquely identified and matched to previously collected data for real-time validation/QC will be considered along with other possible research topics. Similar ideas were considered when defining the 2020 Census research agenda, but our focus was devoted to the efforts to use administrative records for reducing the NRFU workload.

2.k. COMPASS and MOJO: It is suggested there be interaction between Compass and Mojo when it comes to new building data.

Census Response:

Thank you for the suggestion. The Census Bureau will explore ways to improve the interchange of information between these systems for new units/addresses encountered during the Nonresponse Followup Operation.

2.I. BARCA: In the discussions of using satellite imagery to do more in-office address canvassing, the BARCA demo is lagging behind what current technology can do. Specifically, BARCA is relying on humans to visually compare images block-by-block. This is better than people walking all the blocks, but there are several software systems that can compare images to characterize the changes.

Census Response:

The Census Bureau recognizes the review of imagery could be more efficient via an automated process; we are currently investigating automated imagery review capabilities.

3. National Content Test Study Plan

3.a. No clear and concise objective was stated in the NCT Study regarding the ultimate objective, other than an implicit design goal of maximal accuracy in all OMB categories. Even in terms of accuracy, this may conflict with maximizing the percentage of respondent answers that correspond to the respondent's own perceived self-identity (e.g. someone may self-identify as Latino, whereas the OMB designation may be black Colombian). The lack of a clear objective could complicate the ability of the group that will recommend the format of the racial/ethnicity questions to the OMB to reach consensus, particularly since the data that will be presented to the recommending group is likely to be conflicting. In particular, while the NCT is likely to eliminate most questionnaire designs as being clearly inferior, it is also quite likely that there will be multiple designs that will be competitive with one another, with no single design dominating all the others in all respects. For example, some such designs may tend to overcount in some groups and undercount in others, and other designs may offer reverse overcounts/undercounts. The challenge in reaching a final consensus may be compounded by the enormous number of different undercount/overcount "metrics" that the NCT Study will produce. Furthermore, the Census may wish to consider adding a rationale to the lead-in to these questions for why the government requests such information (as Canada does).

Census Response:

The overall objective for the 2015 National Content Test (NCT) race and ethnic research – as discussed in the NCT Study Plan (page 7) – was to test alternative versions of the race and ethnic questions to obtain information and maintain or improve upon the 2010 Census Hispanic origin and race question design and data quality. Our goal was to implement research that refines our efforts to address known race and Hispanic origin reporting issues and important racial and ethnic community concerns while improving data in three crucial areas. These three areas include:

- 1) Increasing accuracy and reliability of reporting in the major OMB race and ethnic categories;*
- 2) Collecting detailed data for myriad groups; and 3) Obtaining lower item nonresponse rates.*

To accomplish this, the 2015 NCT research will evaluate and compare different question designs for race and ethnicity. This research provides the primary opportunity to compare different decennial content questions prior to making final decisions about the content for the 2020 Census. The 2015 NCT also presents the critical opportunity to compare the success of different question designs to determine how they perform in new web-based data collection methods using the Internet, smartphone, and telephone response options.

As discussed in the NCT Study Plan, the test also includes a reinterview to assess the accuracy and reliability of the question alternatives for race and origin, which will enable us to evaluate

key research questions, results, and findings to inform recommendations for the 2020 Census. Specific objectives and goals related to the key dimensions of the NCT research for race and ethnicity are discussed in detail in the NCT Study Plan. For example, with respect to the dimension of testing of question format (page 36), our objective is to determine the best question format for collecting race/ethnic data. Our goal is to maintain or improve the quality of the race and Hispanic origin data by using a combined race and Hispanic origin question. The ways in which this dimension will be evaluated are then discussed in detail, with the presentation of research questions, draft data tables, and decision criteria.

3.b. MENA category: One complicating factor in potentially introducing this classification may be that the current political environment could lead some MENA respondents to prefer to identify in non-MENA categories.

Census Response:

In recent years, we have met with leaders and experts from the Arab American Institute, the Arab American Studies Association, and other leading Middle Eastern and Arab American scholars, activists, and organizations about their request for a separate classification and Middle Eastern or North African (“MENA”) category. We have heard from these leaders and experts that a MENA category is needed in order for the MENA community to respond, be recognized, and be tabulated in the Decennial Census.

An integral part of our work has been extensive outreach with MENA leaders and scholars in preparation for testing a distinct MENA category in the 2015 NCT. It was critical for us to get feedback from the community and stakeholders on the proposed format and usage of the category, and as part of this outreach, the Census Bureau held a “Forum on Ethnic Groups from the Middle East and North Africa” in May 2015. We invited community members and experts to review our plans to test a MENA category, and to provide feedback on several aspects of our data collection efforts.

The 2015 National Content Test (NCT) race and ethnic research is testing the feasibility of employing a “Middle Eastern or North African” (MENA) category. The results from the 2015 NCT research will help inform OMB about the reporting of Middle Eastern responses and North African responses and the ways in which respondents self-identify.

3.c. In the National Content Test how do Brazilians get categorized?

Census Response:

A key component of the race and ethnic research for the 2015 National Content Test (NCT) is to understand how respondents self-identify their race/ethnicity, given different question formats, categories, examples, instructions, and terminology. We frame the presentation of the categories and examples within the current OMB definitions (standards) for race/ethnic groups, and as such, note that the definitions do not fully cover all nationalities and ethnicities around the world. Census Bureau classifications and data tabulations have followed these OMB Standards, and several nationalities such as Brazilian, Belizean, Cape Verdean, and Guyanese, have been classified as part of the “Some Other Race” category. This is also the case for the 2015 NCT.

These countries are not listed in the OMB standards, and their populations are not racially/ethnically homogenous. On the contrary, they have a complex mixture of African, European, indigenous, and Asian origins, among others and respondents self-identify in a multitude of ways. Noting this, the 2015 NCT research will be examining how respondents who report these and other origins, self-identify, and where they specifically report these responses, such as within the Black category, the Hispanic category, the Some Other Race/Ethnicity category, or a combination of categories.

The results from the 2015 NCT research will help inform recommendations about the classification of detailed responses to the race/ethnic questions that are not addressed in the OMB Standards, and these recommendations will be based on what we learn about the ways in which respondents, such as Brazilians, self-identity.

3.d. It would be good to make the data from the content test available to experts on race/ethnicity outside the Bureau for additional analysis. This would be helpful to the Census Bureau.

Census Response:

The 2015 National Content Test (NCT) report will provide a wealth of information and data to document how the research was undertaken, the results for a multitude of research questions, and the discussion and analyses for a series of decision criteria. As we did with the 2010 AQE research results, we plan to thoroughly present and discuss the results with our advisory groups, including the Census Scientific Advisory Committee, as well as myriad stakeholder organizations, race/ethnic scholars, and others to get their critical feedback and thoughts on next steps.

We recognize the profound importance of this research, and we look forward to thoughtful, challenging, and encouraging dialogues with our key advisors, scholars, organizational leaders, and the general public. We recognize there is great interest in this research, and we embrace these dialogues as they help us frame and set a solid foundation for the 2020 Census.

4. Economic Directorate

4.a. Suggestion: We heard very little from the Economic Census directorate this time. As the 2017 survey goes into the field next year, it would be good to get an update at our next meeting. It also would be good to continue to increase the granularity of economic data, such as in GDP with particular emphasis on Spatialization of this data so it would be possible to map and do spatial analytics.

Census Response:

For the next meeting, the Economic Directorate will prepare a session covering preparations for the 2017 Economic Census and Census of Governments. The 2017 Economic Census goes into the field in 2018, as it is collecting data on the 2017 calendar year. The Census of Governments will go into the field in 2017 though, as it is based on government fiscal years ending between July 1, 2016 and June 30, 2017. If there are topics of particular interest, we will address those topics. Otherwise, we will provide a general update on plans, progress, and research results to date.

The Bureau of Economic Analysis (BEA) produces the GDP and is working to increase the granularity of that data. We have jointly acquired two data sets to date (the credit card data and Palantir/First Data), and have met together regularly with these vendors and shared our

successes and challenges with working with the data. Just as with our efforts to increase the geographic granularity of the economic data that we release, we are aware that BEA is looking at increasing the geographic granularity of the GDP data that they produce. Going forward, we anticipate that the two agencies will continue to acquire and work on common data sets, and work more closely on Big Data projects so that we will be looking at common granularity.

5. Working Group on Census Products, the Big Data Center, and the Working Group on Respondent Burden in the ACS

5.a. Recommendation: We recommend the formation of a Data Products Working Group. The group could advise on desirable possible data products. It could also alpha test data products under development. This group might be one place for data from the Census tests to be examined. Also, as interested, a larger number of CSAC members could examine data from the Census tests.

Census Response:

The Census Bureau will consider the possibility of creating a new working group on Data Products.

5.b. Suggestion: We also suggest a session at the next CSAC meeting which “brainstorms” about potential Census data products.

Census Response:

The Census Bureau will consider providing time to the committee for brainstorming potential Census data products during the next in-person committee meeting this fall.

5.c. Suggestion: A presentation on the role and prevalence of proxy interviews would be appreciated.

Census Response:

The Census Bureau will consider providing a presentation on the role and prevalence of proxy interviews on ACS during the next in-person committee meeting this fall.

5.d. Comment: In response to a CSAC question, John Abowd sent a link to the org chart itself, which appears to date from 2011, and to the incumbent heads of each office, which has a March 2016 date. It would be helpful if Sara included the latest versions in materials sent before each meeting.

Census Response:

The CSAC coordinator will continue to provide the agency’s organizational charts (official and operational) as they are updated and prior to each in-person committee meeting.

5.e. Suggestion: There should be meetings with USPS to facilitate collaboration on many fronts, including address data, location points of delivery and routing. USPS are large GIS users and they expressed interest for more coordination. This is not a technical or financial thing, but an administrative and legal one.

Census Response:

The U.S. Census Bureau (Census Bureau) is collaborating with the United States Postal Service (USPS) on a number of activities. Since 2000, the Census Bureau has received the Delivery Sequence File (DSF) every six months from the USPS. The DSF is a USPS data file that contains a

list of all known living quarters in the United States and Puerto Rico; we also receive a monthly address list. The Census Bureau receives numerous other data files from the USPS including an Undeliverable-As-Addressed data file that is being used in 2020 census testing research to assist in the identification and removal of vacant housing units to help reduce the Nonresponse Follow-up universe.

The Census Bureau and the USPS are also collaborating on a number of projects and research, including:

- A Census Bureau proposal to link physical addresses from E911 to PO Box addresses as a service. The USPS has infrastructure in place for this service but not the database of PO Boxes;*
- Automating data file exchanges. Currently the Census Bureau is manually downloading USPS files for many different projects. The Census would like to automate these downloads. Members of the Census Bureau and USPS IT staff have met on two occasions to identify requirements each agency will need to establish an inter-agency IT security agreement. This is an important activity since this includes the tracking needed for outgoing Census questionnaire delivery;*
- Working with the USPS on the Census Bureau's geographic partnership shipping needs and how the USPS can work within the constraints and workflow of the Census Bureau to meet census shipping deadlines for geography programs such as the Boundary and Annexation Survey (BAS) and the Local update of Census Addresses (LUCA) programs;*
- Work with a Census Bureau working group for Puerto Rico to gather information from USPS on the "Every Door Direct or Direct Mail" tool. The working group is documenting how this USPS tool can be utilized by the Census Bureau to save on Nonresponse Followup activities and increase staff recruiting. The Census Bureau is hosting a training session by the USPS on this tool;*
- Identifying ways in which two address data collection projects, the Census Bureau's Geographic Support System (GSS) project and USPS County project can benefit from working collaboratively with county and state governments on address collection. The projects have similar goals and requirements;*
- The development of an Inter-Agency Agreement (IAA) for fingerprinting and onboarding of Census Bureau field staff. This model was used during the 2015 Census Test and will be implemented for the 2017 Census Test;*
- The possibility of the USPS providing access to the geographic representation of their routing data without the address information. This data is very important to the Census Bureau as it will help with the identification of areas with known PO BOX delivery and help with the establish of the correct Type of Enumeration Areas (TEA).*

There are currently two on-going meetings between the Census Bureau and the USPS. The first is a weekly meeting to provide status of on-going projects and research. The second is a quarterly meeting to discuss any issues or concerns with the DSF and/or other data files received by the

Census Bureau from the USPS. When needed other periodic meetings between the Census Bureau and the USPS occur.

5.f. BigData: We were interested in the progress with the BigData center. However, it is still unclear what is a “big data project” and what is a “normal sized data project.” It is suggested that there be a consortium of organizations, facilitated by the center or one of the university partners. This consortium should include universities, owners of big data and technology companies. Passive data collection was discussed, but it is unclear how this is currently used/defined at census. Are there any considerations of doing passive data collection around the online response for the Decennial Census, such as by using web logs.

Census Response:

The Economic Directorate is examining the use of passive data collection. We are currently researching the use of publicly available building permit APIs for use in the Survey of Construction. We have also built a web crawler that can scrape data from public websites. Our research was on state and local government websites, which surveys of local governments can use. We do not believe that the web crawler is limited to public sector sites or surveys. The Economic Directorate is also researching the use of regularly supplied files from companies or from third-party sources that our Economic surveys can use to complete items on our questionnaires, thus substantially reducing respondent burden. Likewise, we plan to look at collaborating with leading accounting software companies to determine if Census reports can be generated from the software.

The Big Data Center ramp up is underway and on schedule, per our FY 16 priorities. The Center's new chief, Dr. Sudip Bhattacharjee, will begin work at the Census Bureau on August 22, 2016. An economist to support the Big Data Center activities began working on June 27, 2016; and a data scientist has accepted the Census Bureau's offer and will begin work on August 22, 2016. These are fully dedicated resources who will supplement a project manager and data scientist who started in October 2015. In FY 17, we anticipate hiring two data scientists through the American Association for the Advancement of Science. Additionally, to further infuse our organization with the necessary, new skillsets, we are now working with the Human Resources Division and the Acquisition Division to identify and leverage data scientist recruitment and hiring services, such as those provided by the Insight Data Science Fellows Program or other comparable programs.

The decennial census operation will collect paradata as part of internet data collection. However, these data will not be used to enhance any existing responses, or to impute any unanswered responses. The paradata are derived from the respondents experience while engaged with the decennial census survey application and these data will be analyzed and used only to inform internet data collection methods and decennial questionnaire content in future tests and censuses. For a number of security and privacy reasons, the Census Bureau does not intend to gather any online or other consumer data about a potential respondents' of the decennial census without consent.

5.g. ACS Respondent Burden Working Group: Even though this was an early report from a working group, CSAC had several concerns about the topics the working group is addressing:

5.g.1. It is not clear what the definition of respondent burden is or what respondents actually find burdensome.

Census Response:

The Census Bureau has worked with the National Academies of Sciences, Engineering, and Medicine's Committee on National Statistics (CNSTAT) in moving toward reduction of burden for the ACS by convening an expert steering committee and conducting a public workshop. The workshop's focus was to shed light on ACS burden concerns and challenges and opportunities to address them through public discussion among a wide range of stakeholders.

One concern raised at the public workshop was how will we measure the respondents' perceived burden of completing the ACS. It was noted that simply reducing the length, in minutes, of the questionnaire may not change the perceived burden if respondents actually view the burden as stemming from the nature of the questions, intrusiveness of government surveys in general, etc. Thus, the suggestion coming out of the workshop was that developing a better metric for perceived burden should be an important first step before introducing major changes to the survey aimed at reducing that burden. Based on the expert meeting discussions regarding this topic, Census Bureau staff are considering ways to measure and evaluate perceived respondent burden. This will continue to be a topic for the remaining expert meetings following the workshop.

5.g.2. There was considerable uncertainty about what a matrix-sampling design would look like for the ACS.

Census Response:

The Census Bureau is not proposing a specific matrix sampling design at this point. A Census team developed an initial feasibility report of various applications of matrix sampling and other methods to reduce the number of questions asked of individual households in the report "Reducing Respondent Burden in the ACS: A Feasibility Assessment of Methods to Ask Survey Questions Less Frequently or of Fewer Respondents" (See <https://www.census.gov/programs-surveys/acs/operations-and-administration/2015-16-survey-enhancements/reducing-respondent-burden.html>).

This report considered four potential methods including 1.) periodic inclusion of questions, 2.) subsampling, 3.) matrix sampling, and 4.) administrative records hybrid. Rather than providing a detailed proposed design under each of these options, it examined each of these as general approaches, and assessed each approach according to a set of factors that demonstrate the feasibility and impact of the method, and made recommendations for next steps and additional research.

The report found that possible designs for matrix sampling would have potentially costly impacts on survey operations and the accuracy and richness of survey estimates. Therefore, the Census Bureau is seeking input that may help to develop research into efficient and effective designs for matrix sampling. We are especially interested in hearing whether the CSAC Working Group has specific ideas for matrix sampling designs that would reduce any potentially negative impacts on

accuracy and richness of survey estimates, or whether the CSAC Working Group would advise against matrix sampling approaches given the potential negative impacts.

5.g.3. In the move to administrative records, it is unclear when they will be used as response variables and when they will be used to help figure out what to do. When administrative records are used, there should be a validation on a subsample.

Census Response:

The Census Bureau is exploring the use of data sources to replace ACS questions. We are currently identifying and acquiring external data sources, matching them to ACS data, and evaluating the coverage and quality of each data source and the resulting matching. Using an agile approach and coordinating adaptation across program areas, including the 2020 Census Program, the Census Bureau will work to develop strategies for obtaining the desired records, and resolving any policy issues associated with their use. This research will identify matching issues, and the challenges associated with securing external data for all cases. The research will compare distributions between ACS data and external data sources for each topic. Researchers will also document measurement issues, such as definitional differences and reference period alignment, as well as the expected stability of data elements (i.e., whether we should expect changes to external data sources over time).

This research is intended to be a first look at the various topics to document the coverage, quality, and availability of external data sources for potential ACS integration. This research will enable ACS to evaluate the potential of the replacement data sources, identify challenges, and provide direction for further research.

A prioritized list of ACS question topics to be studied has been developed, based on the availability of data sources and likelihood of successful matching. We have released several preliminary feasibility assessments as part of this effort: Telephone Service (report released in September 2015); Year Built (report released in November 2015); Condominium Status (report released in November 2015); and Income (report released in March 2016). These reports are available on the ACS website. We are currently working on the Residence One Year Ago feasibility assessment. Of the items we have assessed, the year built and income data are the most promising. Our plan is to develop a research agenda to determine how to best use administrative data in the ACS. Census is consulting with external experts and stakeholders on this project.

As we identify candidate data sources that have acceptable coverage and quality, we will plan to use a 5-year ACS data set to create a simulated 5-year product using administrative records that would be compared to ACS estimates. We acknowledge that in some cases we cannot assume that the ACS data reflect "truth" when there are differences between the ACS estimate and a simulated administrative records estimate, given that each data source has its own sources of error. Early research would focus on proving this option in operationally, assessing the data quality, and using it to understand changes in published data that may come from the use of administrative data. This would potentially be followed by a field pilot where questions are skipped in automated modes if data are available for that address. The specific design of that field pilot has yet to be determined, but could include validation on a subsample.

5.g.4. It was also noted that the business community comprises an important set of users of the ACS PUMS data.

Census Response:

The Census Bureau agrees that the business community comprises an important segment of users interested in our ACS PUMS data. That is why we invest time and effort in reaching this community through webinars, conferences, presentations (e.g., Big Data Innovation Summit, America's Small Business Development Centers Annual Conference, Select USA Investment Summit). During the development of the Data Products Redesign Group (DPRG) initiative, the Census Bureau actively recruited from groups we knew had a particular interest in data products such as the PUMS files. The business community was one of these groups and we are happy to include representatives of Nielsen, Five Thirty Eight, and a private consulting firm among the DPRG members.

5.g.5. An update on the status and activities of the ACS Users' Group would be helpful.

Census Response:

The Population Reference Bureau (PRB) launched the new ACS Online Community on April 25 to facilitate communication among ACS data users. This online community provides a discussion site where people can share messages, materials, and announcements related to ACS data products, methods, events, and other ACS data issues.

The new site has a responsive design, so that users can view the site on laptops, smartphones, etc. Other new features include the following: navigation tabs at the top of the page, parts of the page can be made public or private, a twitter feed, videos and hyperlinks can be added to discussions, and a site feedback form that can be submitted.

This site, like the previous site, includes a variety of groups that members can join, for example, Public Use Data Sample Files, Summary Files, Application Programming Language (API), Data for Small Geographic Areas and Population Subgroups, Mapping ACS Data/GIS applications, and Measuring Trends Over Time.

There are currently over 1,600 members. About 70 members of the Online Community participated in a webinar on April 28, 2016 that provided an overview of the new site. A link to the recording of the webinar is posted in the Online Community.

<https://acsdatacommunity.prb.org/>

Last year, the Census Bureau contracted with the Population Reference Bureau to form and manage an informal (non-advisory) Data Products Redesign Group (DPRG). The group includes members from a broad range of ACS data user communities with different skill or experience levels. It was formed to ensure that future products are customer focused and to assist the Census Bureau in understanding the needs of our data users. The DPRG will provide an initial source of feedback on dissemination channels that best meet our data users' needs. Feedback from the DPRG will include providing feedback on the overall ACS data products plan, ideas for new types of data products, reaction to specific product additions/deletions, and early testing of new dissemination channels or tools.

In addition, the DPRG is testing alpha versions of the Center for Enterprise Dissemination Services and Consumer Innovation (CEDSCI) platform. The CEDSCI will consist of the services and infrastructure to handle data dissemination for all Census Bureau surveys/censuses.