

An Overview of 2020 Census Design

Robin A. Pennington¹, Michael Bentley, Gina Walejko
U.S. Census Bureau, 4600 Silver Hill Road, Suitland, MD 20746
U.S. Census Bureau, 4600 Silver Hill Road, Suitland, MD 20746
U.S. Census Bureau, 4600 Silver Hill Road, Suitland, MD 20746

Abstract: *The 2020 Census Operational Plan defines the operations to be performed in the 2020 Census and describes the function and scope of each operation. Since the 2010 Census, the Census Bureau has been designing and testing technological advances and other innovative methodologies for completing the next census, which are outlined in the operational plan. In this paper, current plans for major redesigned operations will be presented. Of particular note, the 2020 Census will include numerous innovations to maximize self-response, especially through online response. Other strategies to maximize self-response and challenges associated with maximizing self-response for subpopulations will also be discussed. Additional innovations with significant impacts to census design such as in-office development of the census address frame and targeted uses of administrative records to reduce field operation workloads will also be described.*

Keywords: Operational Design, Optimizing Self-Response, Integration of Operations

1. 2020 Census Background

Many improvements to the collection of data, automation of operations, efficiency of fieldwork, and integration of data collected from the operations are planned for the 2020 Census. The 2020 Census Operational Plan (U.S. Census Bureau, 2017), detailing the full plan for the upcoming census, was first released in 2015. This document has been completely updated annually, and version 4.0 of this plan will be released in December 2018. This final version before the 2020 Census will incorporate any lessons learned and design updates that result from the testing in the 2018 End-to-End Census Test. In tandem with the publication of full plans for the census in this version of the 2020 Census Operational Plan, this paper will present some specific details about a number of the primary innovation areas.

First, we will discuss the efforts to maximize self-response, both through partnership and communications efforts and through a refined strategy for contacting households and informing the public about the census. A second major innovation to the design of the census for this decade is the use of selected administrative records to supplement and enhance operations and field results; the latest status of this effort will also be described. Finally, some of the architectural diagrams related to integration of these redesigned operations help ensure that tasks will be completed and will run smoothly. Some of the key

¹ *The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.*

artifacts that are of use for clarifying and explaining the planned implementation for the 2020 Census will be included here, as well as in the fourth version of the 2020 Census Operational Plan.

The goal of the 2020 Census' "Optimizing Self-Response" innovation area is to make it as easy and efficient as possible for people to respond to the 2020 Census by offering response options through the internet and telephone, in addition to a paper questionnaire. Self-response reduces the need to conduct in-person follow-up operations to complete the enumeration. To that end, the Census Bureau will motivate people to respond, as well as make it easy for people to respond, from any location at any time, even if they do not have the identifier provided to them by the Census Bureau. The internet response option is intended to make response easier than ever before, at the same time improving the speed of the data collection and the quality of the household data. While this response mode promises many benefits, the Census Bureau fully recognizes the limitations of imposing this mode of response on the entire U.S. household population. The Census Bureau has been experimenting all decade with contact strategies to inform the population of the availability of this option while leaving prior response modes available.

Internet response marks a significant methodological change for this decennial census. From 1970 to 2010, the primary method of response prior to Nonresponse Followup (NRFU) was to mail paper questionnaires and ask the public to complete and return them by mail. Between each census, the Census Bureau tweaked the mailing design in order to increase response rates, such as adding a single reminder postcard and a second paper questionnaire, but the general design remained little changed. As the Census Bureau modernizes the 2020 Census, for much of the country the primary means of responding will be to go online. The 2020 Census will be joining many other surveys, both at the Census Bureau, such as the American Community Survey, and in other government and private sector surveys, that allow for online responses.

2. Optimizing Self-Response – Research and Strategies

Making a big methodological change requires significant testing and refinement. The 2012 National Census Test was the first test after the 2010 Census and tested mail contact strategies. This research included several variations on the order, format, and content of the mailing materials. The Census Bureau learned that an initial contact to encourage online response followed by two reminders and then, finally, a paper questionnaire was very promising, with an increase in response rates of more than 4 percentage points compared to the control design. The Census Bureau now refers to this general approach as **Internet First**, and it has served as the control design in tests ever since. The Census Bureau also found that an advance letter with no action required of the respondent did not improve response rates and that telephone assistance was needed for respondents who could not go online.

The 2014 Census Test, a site test in Washington, D.C. and Montgomery County, Maryland, examined alternative contact strategies including email, automated voice messages, and the ability of respondents to "pre-register" to be reminded to fill out the form. The test concluded that pre-registration did not have a positive impact on response rates.

In 2015, several tests occurred. The 2015 Optimizing Self-Response site test in Savannah, Georgia examined how advertising, outreach, and promotion can engage and motivate respondents for action. It also tested the ability for households to respond without an ID.

One notable finding was that housing units that were not selected to receive a mailing still responded using the internet without an ID as a result of advertising and promotional efforts. The 2015 National Content Test, a nationally representative sample of 1.2 million addresses, primarily tested census content matters. It also tested mail material strategies, including sending a paper questionnaire up front to geographies with low internet penetration, known as **Internet Choice**, while pushing other geographies to respond online by withholding the paper questionnaire until the fourth mailing, known as **Internet First**. One major finding of the 2015 National Content Test was that adding a fifth mailing, an additional reminder after the paper questionnaire, further increased self-response rates. Also, the Internet Choice strategy was shown to be the most effective in low response areas. In addition, not sending a paper questionnaire to nonrespondents resulted in the lowest self-response rates across all groups.

The 2016 Census Test was conducted in parts of Los Angeles, California and Houston, Texas, and it tested methods to increase support for non-English speakers by variations in the mail materials including a multilingual brochure and an FAQ insert. These were both shown to be effective at increasing the number of responses in non-English languages (Spanish, Chinese, or Korean).

The 2017 Census Test was a nationwide self-response test looking at the feasibility of collecting tribal enrollment information. Self-response was tracked, but there were no experimental findings related to optimizing self-response that impacted the 2020 Census design. The 2018 End-to-End Census Test in Providence County, Rhode Island was an end-to-end test of systems and operations planned for the 2020 Census. The test concluded in August 2018. Further refinement of estimates and procedures will occur based on the results of this test.

In each of the tests mentioned above, the Census Bureau tracked the response rates by mode. Table 1 shows response rates by mode compared to 2010 Census response rates. Overall response rates are lower in tests than in censuses, but we see high rates of internet response within the overall self-response.

The culmination of test findings thus far has led to the contact strategy shown in Figure 1. In the **Internet Choice** contact strategy, questionnaires are still sent to the household at the initial contact. The **Internet First** contact strategy will undergo four waves of mailings, known as “cohorts”, occurring over a week to spread out the volume of calls to Census Questionnaire Assistance and also to better control the number of people accessing the online instrument each day.

Table 1: Response Rates for Decennial Census Tests Occurring between 2012 and 2018	Internet	Phone	Mail	Total	2010 Census
2012 National Census Test (50 states and D.C.)	36.5	7.4	13.0	56.9	66.5
2014 Census Test (Montgomery County and Washington D.C.)*	50.6	5.2	10.2	65.9	72.8
2015 Census Test (Maricopa County)*	39.7	6.4	8.8	54.9	63.8
2015 Optimizing Self-Response Test (Savannah)	33.4	5.6	8.4	47.5	56.5
2015 National Content Test (50 states and D.C., average of all panels)	35.6	6.1	10.1	51.9	66.5
2016 Census Test (Harris County)*	28.2	1.9	9.2	39.3	60.5
2016 Census Test (Los Angeles County)*	32.4	2.6	17.4	52.5	71.6
2017 National Census Test (50 states and D.C.)	31.7	2.4	16.2	50.3	66.5
2018 Census Test (Providence County) (Preliminary results)	32.0	3.9	16.4	52.3	64.7
*Test conducted in a portion of the selected area Source: Response data and sample operational data for each individual test.					

Panel	Cohort	Mailing 1 <i>Letter (Internet First) or Letter + Questionnaire (Internet Choice)</i>	Mailing 2 <i>Letter</i>	Mailing 3* <i>Postcard</i>	Mailing 4* <i>Letter + Questionnaire</i>	Mailing 5* <i>“It’s not too late” Postcard</i>
Internet First	1	March 12, 2020	March 16, 2020	March 26, 2020	April 8, 2020	April 20, 2020
	2	March 13, 2020	March 17, 2020	March 27, 2020	April 9, 2020	April 20, 2020
	3	March 19, 2020	March 23, 2020	April 2, 2020	April 15, 2020	April 27, 2020
	4	March 20, 2020	March 24, 2020	April 3, 2020	April 16, 2020	April 27, 2020
Internet Choice	N/A	March 13, 2020	March 17, 2020	March 27, 2020	April 9, 2020	April 20, 2020
*Mailings targeted only to nonresponding households						

Figure 1: Planned 2020 Census Mail Contact Strategies (Internet First and Internet Choice)

We have just described our research into increasing self-response through changes in mail material strategies, but reaching potential respondents through communications and outreach efforts is also a key part of the “Optimizing Self Response” innovation area. The goal of the Census Bureau’s Integrated Partnership and Communications Program is to engage and motivate people to self-respond, preferably by internet, and to raise awareness that encourages response throughout the entire 2020 Census, including during NRFU. As shown in Figure 2, this decennial census operation has several components, of which five are considered “major” components: website and mobile, media relations, paid advertising, statistics in schools, and social media. Partnership efforts are considered central to each of these components. The integration of these components will ensure consistent messaging throughout the 2020 Census.

All components of the Integrated Partnership and Communications Program will be driven by research, particularly the paid advertising component. The Census Bureau will use a census tract’s propensity to self-respond, demographic data, media consumption information, and attitudinal information from the 2020 Census Barriers, Attitudes, and Motivators Study (CBAMS) Survey to segment tracts into audiences. The Census Bureau will use this audience segmentation to design advertisements and plan media buying. After advertising is delivered, the Census Bureau will address issues that arise during the 2020 Census and coordinate with field and partnership teams to prioritize audiences and align messages. (See Figure 3 for basic steps related to research behind the Integrated Partnership and Communications Program.)



Figure 2. Integrated Partnership and Communications Program Components

Research and Communications Integration



Figure 3. Eight Major Steps in Research Related to Integrated Partnership and Communications Program

3. Utilizing Administrative Records and Third-Party Data

Another major innovation planned for the 2020 Census is the use of administrative records or third-party data for the following purposes: determining the status (i.e. vacant, occupied, or nonexistent) of a housing unit listed on the Census Bureau's address list; imputing missing demographic characteristics from census returns that otherwise satisfy form completion criteria; filling in household data for households that have not responded after all of the attempts in the implemented optimized contact strategy, when the existing data are deemed to be of sufficiently high quality.

While many papers this decade have described these planned uses of administrative records (Keller & Konicki, 2016; Morris et al, 2016; U.S. Census Bureau, 2017), we focus here on high-level descriptions of the methodology. In particular, we describe in sufficient detail the uses that will appear in the integration diagrams presented at the end of this paper. Figure 4 shows the various decision points where administrative records can be used to resolve an enumeration case. In particular, administrative records will be used at the start of enumeration operations to determine an initial status of units. Units flagged as vacant or nonexistent will be sent an additional mailing to validate the status. Units flagged as occupied but that have not self-responded will be included in NRFU. Enumeration can occur in NRFU—or during the same timeframe through self-response—but after some number of NRFU enumeration attempts, administrative records of sufficient quality will be used for the enumeration. Finally, cases that remain unresolved after the maximum number of NRFU attempts could use administrative records data in imputation procedures.

4. Integration of Operations

In order these innovations to work, integration across operations is needed to an even greater extent than in prior censuses. For example, perhaps spurred by the optimized mailing contact strategy or targeted advertising, a household can self-respond late to the census through any of the available modes, and that response will trigger an automatic removal of the case from the NRFU workload. This can occur even after an unsuccessful attempt to contact the household in NRFU has occurred. Furthermore, tracing data from the United States Postal Service indicating that the response is in the mail may be sufficient to remove the case from further contact attempts. The usage of administrative records and third-party data for the enumeration of an occupied household in place of a household response requires careful coordination of file vintages, file deliveries, modeling results, and determination and delivery of an enumeration status across multiple systems.

Various documents and diagrams are used to enable such integration. Systems architecture diagrams show the planned interactions of all the systems that will be implemented and integrated, while operation interactions are documented within the Detailed Operational Plans (DOPs) that are published for nearly every operation within the census. These DOPs show Business Process Models, which are a graphical representation of business processes or workflows, and input/output diagrams, which show the inputs needed to perform an operation, the outputs from an operation, the stakeholders and owners, and the systems required to perform the operation. Input/output diagrams can be checked against each other to determine if every input coming to one operation is shown as an output from the other operation, and, similarly, if every output from one operation is shown as an input to the other operation. Additionally, an overlay is created on the Business Process Models that shows the groupings of tasks at a higher level, in what is known as an Activity Tree.

Activity Trees across all the operations can be grouped together, resulting in an activity diagram that displays all the functions within the census and provides additional insight into the design. For example, large branches on the tree that have few tasks may not be designed most efficiently, while small branches that have a large number of tasks may not be defined at the right level. A high-level Activity Tree for the 2020 Census is shown in Figure 5, with a break-out into some of the branches for the NRFU operation within the Data Collection operations. Other diagrams showing operational interactions that are of use in understanding the design of the 2020 Census are the Top-Level End-to-End Operational Flow, the Production Timeline Analysis, and finally the Integrated Operations Diagrams that appear in most DOPs, as well as in the multiple DOPs for the operations that require integration. Both the Top-Level End-to-End Operational Flow and the Production Timeline Analysis will be included in v4.0 of the 2020 Census Operational Plan.

In summary, while 2020 Census planning has innovated processes and specific operations, for the most effective census, these operations will need to integrate their efforts with other operations in meaningful ways. In this paper we have described innovations associated with the 2020 Census mailing strategy and briefly discussed the integration of this strategy with the internet response mode. Our Integrated Partnership and Communications will also integrate with both self-response and the NRFU operation for messaging to the public, while administrative records will be integrated into both self-response and NRFU in order to resolve cases where no response has been received. In particular, this includes both when a response cannot be received because a housing unit on the Census Bureau's address list is vacant or nonexistent as well as when a housing unit is believed to be occupied. We have also provided diagrams that describe integration of the related operations.

Acknowledgements

Integration artifacts have been created through work with the MITRE Corporation.

References

Keller, A., & Konicki, S. (2016). Using 2010 Census Coverage Measurement Results to Better Understand Possible Administrative Records Incorporation in the Decennial Census. In JSM Proceedings, Survey Research Methods Section. Chicago, IL: American Statistical Association. 701-710.

Morris, D., Keller, A., Clark, B. (2016). An approach for using administrative records to reduce contacts in the 2020 Decennial Census. *Statistical Journal of the IAOS*, vol. 32, no. 2, pp. 177-188.

U.S. Census Bureau. (2017). 2020 Census Operational Plan. Washington DC: Census Bureau. Available at: <https://www2.census.gov/programs-surveys/decennial/2020/program-management/planning-docs/2020-oper-plan3.pdf> (accessed September 2018).

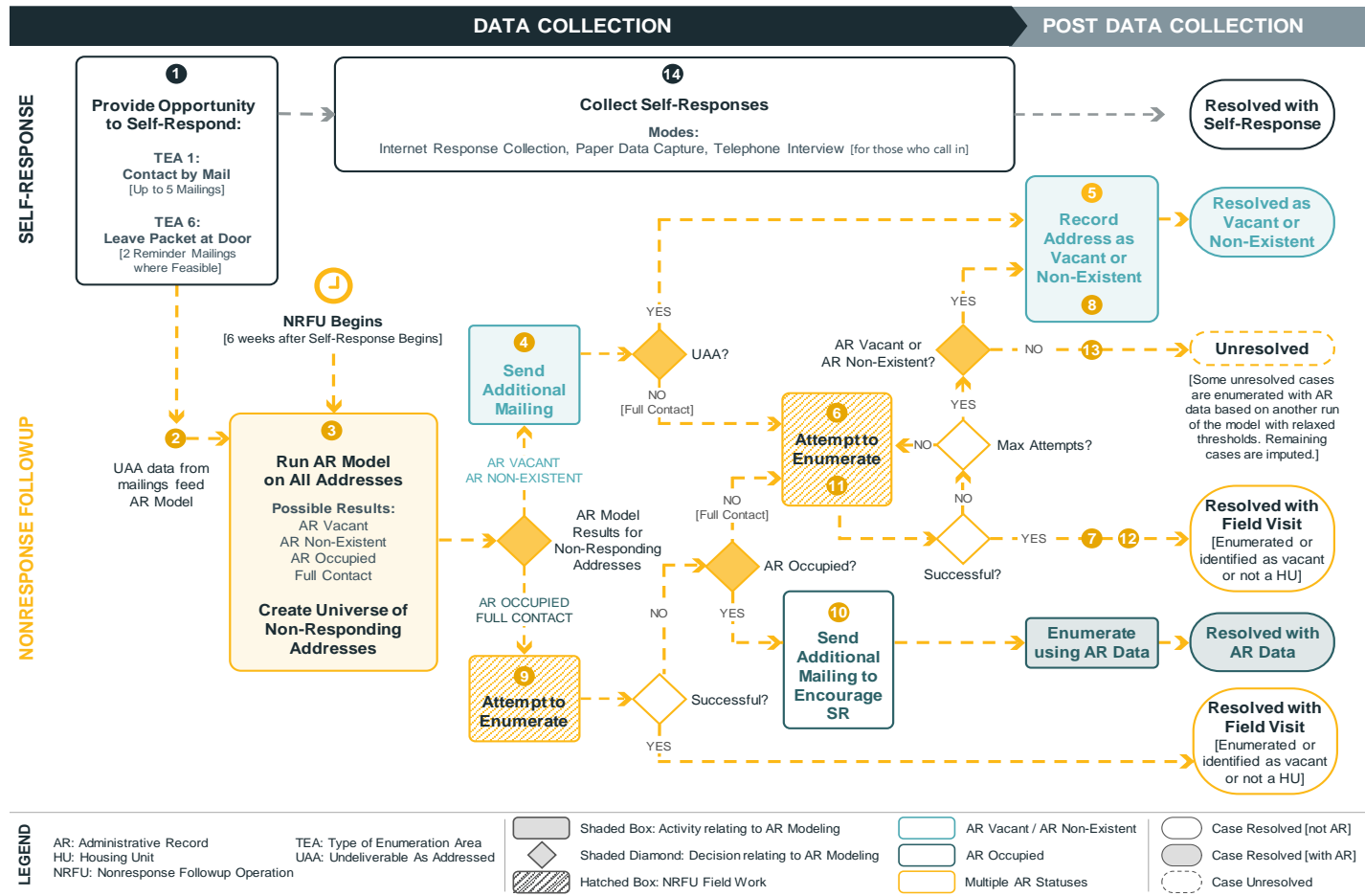
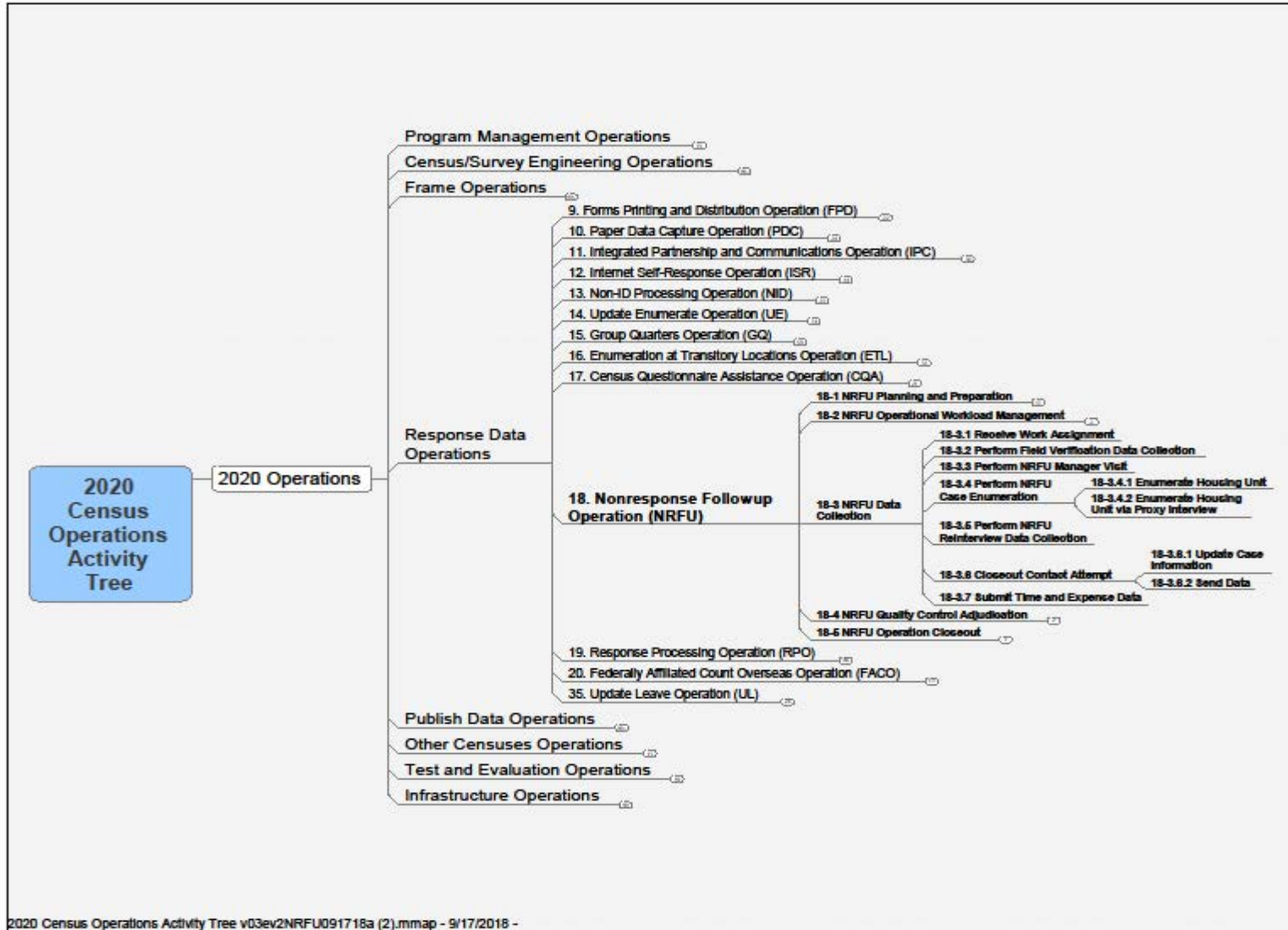


Figure 4. Administrative Records Decision Points



2020 Census Operations Activity Tree v03ev2NRFU091718a (2).mmap - 9/17/2018 -

Figure 5: 2020 Census Operations Activity Tree and Branch