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

September 7, 2017

2017 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT MEMORANDUM SERIES # ACS17-RER-09

MEMORANDUM FOR	Victoria Velkoff Chief, American Community Survey Office
From:	David Waddington Chief, Social, Economic, and Housing Statistics Division (SEHSD)
Prepared by:	Jamie Lewis Social, Economic, and Housing Statistics Division (SEHSD)
Subject:	2016 American Community Survey Content Test Evaluation Report: Computer and Internet Use

Attached is the final report for the 2016 American Community Survey (ACS) Content Test for Computer and Internet Use. This report describes the results of the test for the revised versions of the Types of Computer, Internet Access, and Internet Subscription questions.

If you have any questions about this report, please contact Kurt Bauman at 301-763-6171 or Jamie Lewis at 301-763-4535.

Attachment

cc:
Agnes Kee (ACSO)
Jennifer Ortman (ACSO)
Edward Porter (CSRM)
Dorothy Barth (DSSD)
Patrick Cantwell (DSSD)
Asaph Young Chun (DSSD)
Elizabeth Poehler (DSSD)
Anthony Tersine (DSSD)
Kurt Bauman (SEHSD)
Nicole Scanniello (SEHSD)

Intentionally Blank

September 7, 2017

2016 American Community Survey Content Test Evaluation Report: Computer and Internet Use

FINAL REPORT

Jamie M. Lewis
Social, Economic, and Housing Statistics Division

Dorothy A. Barth
Decennial Statistical Studies Division

Intentionally Blank

TABLE OF CONTENTS

List of Tables

List of Figures

EXECUTIVE SUMMARY

Overview

From February to June of 2016, the U.S. Census Bureau conducted the 2016 American Community Survey (ACS) Content Test, a field test of new and revised content. The primary objective was to test whether changes to question wording, response categories, and definitions of underlying constructs improve the quality of data collected. Both new and revised versions of existing questions were tested to determine if they could provide data of sufficient quality compared to a control version as measured by a series of metrics including item missing data rates, response distributions, and response error. The results of this test will be used to help determine the future ACS content and to assess the expected data quality of revised questions and new questions added to the ACS.

The 2016 ACS Content Test consisted of a nationally representative sample of 70,000 residential addresses in the United States, independent of the production ACS sample. The sample universe did not include group quarters, nor did it include housing units in Alaska, Hawaii, or Puerto Rico. The test was a split-panel experiment with one-half of the addresses assigned to the control treatment and the other half assigned to the test treatment. As in production ACS, the data collection consisted of three main data collection operations: 1) a six-week mailout period, during which the majority of self-response via internet and mailback were received; 2) a one-month Computer-Assisted Telephone Interview period for nonresponse follow-up; and 3) a one-month Computer-Assisted Personal Interview period for a sample of the remaining nonresponse. For housing units that completed the original Content Test interview, a Content Follow-Up telephone reinterview was conducted to measure response error.

Computer and Internet Use

This report discusses the topic of Computer and Internet Use, which was first introduced to the ACS in 2013. Because of the rapid change in technology and terminology, it was evident that the questions regarding this topic needed to be revised. Specific concerns included the relatively low percentage of handheld-owning households reporting an internet subscription or a mobile broadband subscription.

For the question on computer usage, the number of response categories increased from three to four (with a new category for tablet computers) and the wording of each category was revised for clarity, such as replacing “Handheld computer” with “Smartphone.” The wording was revised for the internet access question to address the rapid change in how people access the internet and the terminology we use to describe internet access, asking about payment (rather than subscription) to a cell phone company in addition to an internet service provider.

For the internet subscription question, the number of response options was reduced from seven to five (dropping “DSL,” “Cable modem,” and “Fiber-optic” as separate categories), wording was revised for clarity, and the phrase “Mobile broadband plan” was replaced with “Cellular data plan.” The options were also presented in a different order. Although the test versions of the computer and internet use questions were implemented in 2016 production ACS (Reichert,

2015), the topic was included in the Content Test in order to conduct analysis to validate the early implementation decision.

Research Questions and Results

This research was guided by several research questions concerning missing data rates, differences in the reports of computer usage and internet subscriptions by treatment, and response error. Although not part of the key research, comparisons were also made to benchmark estimates from the Current Population Survey (CPS) Computer and Internet Use Supplement and surveys conducted by the Pew Research Center. Research questions, methodology, and results on benchmarks can be found in Appendix C.

Item Missing Data Rates

Results indicate that the item missing data rates are not significantly different between treatments for the types of computers question as a whole, as well as for the individual computer categories. For the internet access question, the item missing data rate is significantly lower in the test treatment than in the control treatment, indicating that the test version of the question performed better. For mail responses, any response that indicates more than one type of internet access (a response with more than one box being marked) is considered "missing" data. The rate at which this type of response occurs for the internet access question is not significantly different between treatments, indicating that the changes to the question did not affect this indicator. Finally, for the internet subscription type question overall, the item missing data rates are not significantly different between treatments. Of the five categorical comparisons made, the only item missing data rate that shows a significantly lower value in the test treatment compared with the control treatment is the rate for the "Cellular data plan" category.

Response Proportions

Findings for the types of computers question reveal that the proportion of "Yes" responses for the "Desktop or laptop" category is lower in the test treatment than in the control treatment. A possible explanation is the introduction of a separate "Tablet" category to the test version of the question. In the absence of this category, some control respondents owning tablets (but not desktops or laptops) may have marked the category for "Desktop, laptop, netbook, or notebook computer." A larger proportion of test households reported owning or using a smartphone or tablet, compared with the share of control households reporting a handheld computer. A smaller proportion of households in the test treatment indicated that they owned or used some other computer, compared with the control treatment.

Regarding the question on internet access, among all households overall and households with a smartphone or tablet, the proportion reporting an internet subscription is higher in the test treatment than in the control treatment. Similarly, reporting of no internet subscription is lower among test households overall.

For the final item on internet subscription type, reports of mobile broadband are strikingly higher in the test treatment (about 80 percent) than in the control treatment (about 40 percent), whether

looking at households overall or focusing on households with a smartphone or tablet. The proportion of households reporting a broadband service such as DSL, cable, or fiber-optic is lower in the test treatment than in the control treatment. While the difference is significant, the magnitude is fairly small and the results are close to what we were expecting. The difference likely reflects the number of categories measuring this type of service. Respondents had three categories of this type in the control version of the question, but only a single category in the test version. There is no significant difference in the share of households reporting a dial-up subscription, satellite internet service, or some other service in the test versus control treatments.

Response Error

The test version of the types of computers question is more reliable than the control version for the categories of smartphone and tablet use and use of some other computer. There were no significant differences between test and control for the other computer categories. For the internet access question, the inconsistency in reports of access with a subscription and access without a subscription is lower in the test treatment than in the control treatment. No other significant differences between treatments were detected for the reliability metrics. For the internet subscription type question, the cellular data plan category in test has greater response reliability than the mobile broadband category in control. The high speed and satellite internet categories in test did not perform as well as control for one of the response reliability metrics. There were no significant differences between control and test for the remaining internet subscription categories.

Conclusion

Overall, results indicate that data quality improved when using the revised questions. All of the key research criteria for the internet access question were met, and four of five key research criteria were met for both the types of computers and internet subscription type questions. In each case, the key criterion not met was of lowest priority.

Item missing data rates in the test treatment were not significantly different from those in the control treatment across the board. Results for the response proportions analysis, in general, were as expected. Particularly noteworthy is the substantial increase in the share of households reporting a cellular data plan in the test treatment versus a mobile broadband plan in the control treatment. Whether looking at all households or specifically at households with a smartphone or tablet (handheld in control), the test proportion is about double the control proportion. Finally, although the reliability of the high speed and satellite internet categories was better for the control version, for the most part the test version of the Computer and Internet Use questions was more reliable or not significantly different from the control version.

Altogether, the 2016 ACS Content Test and analyses presented here validate the decision to implement the revised question wording on the 2016 production ACS. The revised question wording will be reflected in the 2016 ACS data release, scheduled to begin in September 2017.

1 BACKGROUND

From February to June of 2016, the Census Bureau conducted the 2016 American Community Survey (ACS) Content Test, a field test of new and revised content. The primary objective was to test whether changes to question wording, response categories, and definitions of underlying constructs improve the quality of data collected. Both revised versions of existing questions and new questions were tested to determine if they could provide data of sufficient quality compared to a control version as measured by a series of metrics including item missing data rates, response distributions, and response error. The results of this test will be used to help determine the future ACS content and to assess the expected data quality of revised questions and new questions added to the ACS.

The 2016 ACS Content Test included the following topics:

- Relationship
- Race and Hispanic Origin
- Telephone Service
- Computer and Internet Use
- Health Insurance Coverage
- Health Insurance Premium and Subsidy (new questions)
- Journey to Work: Commute Mode
- Journey to Work: Time of Departure for Work
- Number of Weeks Worked
- Class of Worker
- Industry and Occupation
- Retirement, Survivor, and Disability Income

This report discusses the topic questions involving Computer and Internet Use.

1.1 Justification for Inclusion of Computer and Internet Use in the Content Test

The questions collecting information on computer availability, internet access, and internet subscriptions were first introduced on the ACS in 2013. Given the rapid rate at which technology is growing and changing, it became apparent that a revision to these questions was already needed. As an example, prior to January 2016, the question about computer ownership did not specifically ask about tablets, but the rate of tablet ownership has grown dramatically in recent years, with recent estimates indicating that 51 percent of adults in the country own a tablet (Pew Research Center, 2016).

Preliminary data from 2013 showed that the wording of the questions needed to be revised for a variety of reasons. One finding that raised concerns was the somewhat low percentage of handheld-owning households who reported having an internet subscription. For owners of desktops and laptops, internet subscription was reported to be 91.1 percent. For owners of handheld devices, the internet subscription rate was 76.3 percent. We anticipate that the inclusion of the phrase “cell phone company” will encourage handheld-owning households to think about their data plans as internet subscriptions.

Preliminary data also showed that the question wording needed to be revised because of low reports of mobile broadband subscriptions (File & Ryan, 2014). In 2013, for example, only 33 percent of households reported having mobile broadband subscriptions though 64 percent of households reported having a handheld device. Thus, among households with handheld devices and internet subscriptions, only 54 percent reported a mobile broadband subscription. If the estimate of mobile broadband is correct, then half of the households with a handheld device are using them without a data plan. We expect the proportion of households with mobile broadband to increase with the new question wording in the test version.

1.2 Question Development

Initial versions of the new and revised questions were proposed by federal agencies participating in the U.S. Office of Management and Budget (OMB) Interagency Committee for the ACS. The initial proposals contained a justification for each change and described previous testing of the question wording, the expected impact of revisions to the time series and the single-year as well as five-year estimates, and the estimated net impact on respondent burden for the proposed revision.[1] For proposed new questions, the justification also described the need for the new data, whether federal law or regulation required the data for small areas or small population groups, if other data sources were currently available to provide the information (and why any alternate sources were insufficient), how policy needs or emerging data needs would be addressed through the new question, an explanation of why the data were needed with the geographic precision and frequency provided by the ACS, and whether other testing or production surveys had evaluated the use of the proposed questions.

The Census Bureau and the OMB, as well as the Interagency Council on Statistical Policy Subcommittee, reviewed these proposals for the ACS. The OMB determined which proposals moved forward into cognitive testing. After OMB approval of the proposals, topical subcommittees were formed from the OMB Interagency Committee for the ACS, which included all interested federal agencies that use the data from the impacted questions. These subcommittees further refined the specific proposed wording that was cognitively tested.

The Census Bureau contracted with Westat to conduct three rounds of cognitive testing. The results of the first two rounds of cognitive testing informed decisions on specific revisions to the proposed content for the stateside Content Test (Stapleton and Steiger, 2015). In the first round, 208 cognitive interviews were conducted in English and Spanish and in two modes (self-administered on paper and interviewer-administered on paper). In the second round of testing, 120 cognitive interviews were conducted for one version of each of the tested questions, in English and Spanish, using the same modes as in the first round.

A third round of cognitive testing involved only the Puerto Rico Community Survey (PRCS) and Group Quarters (GQ) versions of the questionnaire (Steiger, Anderson, Folz, Leonard, & Stapleton, 2015). Cognitive interviews in Puerto Rico were conducted in Spanish; GQ cognitive

[1] The ACS produces both single and five-year estimates annually. Single year estimates are produced for geographies with populations of 65,000 or more and five-year estimates are produced for all areas down to the block-group level, with no population restriction.

interviews were conducted in English. The third round of cognitive testing was carried out to assess the revised versions of the questions in Spanish and identify any issues with questionnaire wording unique to Puerto Rico and GQ populations.[2] The proposed changes identified through cognitive testing for each question topic were reviewed by the Census Bureau, the corresponding topical subcommittee, and the Interagency Council on Statistical Policy Subcommittee for the ACS. The OMB then provided final overall approval of the proposed wording for field testing.[3]

The development of the computer device question came about as a result of a need to keep up with technological updates and changes in computer terminology. The terminology "netbook" and "notebook" computer were excluded from the test version of the question, because they are outdated terms. We know from recent Pew numbers that around 51 percent of adults in the country own tablets (Pew Research Center, 2016), so it was apparent that the word "tablet" needed to appear in a category. Also, new technology, such as smart watches and Google glasses, are worn and not carried by hand, so the word "handheld" became outdated. As a result, a new category "Tablet or other portable wireless computer" was created. "Smartphone" was created as its own category because they have become so widely used and owned.

The development of the internet access question came about as a result of changes in technology, terminology, and the way people access the internet. During the cognitive testing phase, concerns were raised about the confusing nature of the word "subscription" (Stapleton and Steiger, 2015). Thus, this term was excluded in the final test version of the question. Also, the words "by paying a cell phone company" were added to help respondents realize that their data plans are equivalent to a paid internet service. This addition, in combination with "Smartphone" having its own category as a computer device, should increase the quality of data collected regarding mobile internet access. Also, during cognitive testing, some respondents answered the question thinking about their habits of internet use at home rather than focusing on their actual ability to access the internet at their house. This will undoubtedly become more of an issue as internet access technology grows increasingly mobile. The phrase "access the internet" was changed to "have access to the internet" to more accurately convey the intent of the question.

The question involving types of internet subscriptions was revised in order to address changes in internet use and terminology. The first round of cognitive testing included two categories that used the word *broadband*: "Mobile broadband" and "Broadband (high speed)." Some respondents answered incorrectly because of their misinterpretation of the terms. Several changes were made to the question to address this problem. The phrase "At this house, apartment, or mobile home" was removed from the question and "installed in this household" was added to the end of the "Broadband (high speed)", "Satellite", and "Dial-up" categories. This change should help respondents more easily differentiate between smartphone data plans, which are not tied specifically to a place, and other ways of having access to the internet that are tied to a place. This change also enabled a redesign of the mobile broadband category to put less emphasis on the duplicative use of the term "broadband" by changing "Mobile broadband" to "Cellular data plan." Three categories were collapsed and used as examples to better describe "Broadband (high speed)" internet service.

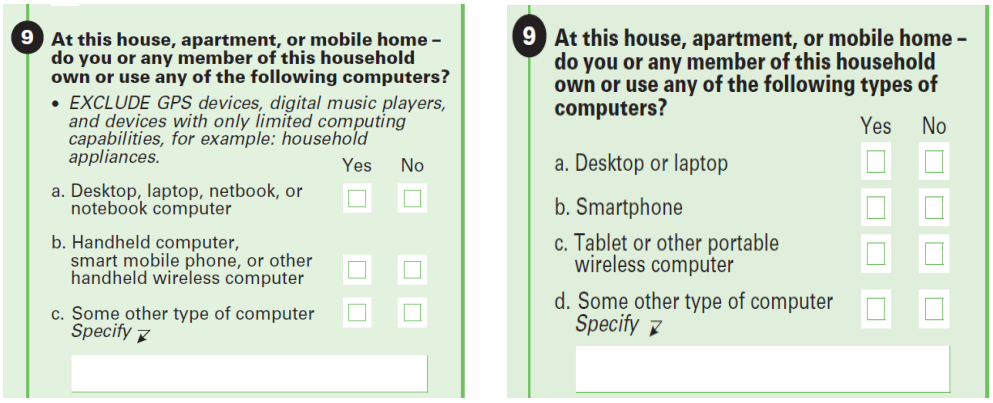
[2] Note that the field testing of the content was not conducted in Puerto Rico or in GQs. See the Methodology section for more information.

[3] A cohabitation question and domestic partnership question were included in cognitive testing but ultimately we decided not to move forward with field testing these questions.

1.3 Question Content

Control and test versions of each question are shown as they appeared on the mail questionnaire. Automated versions of the questionnaire had the same content formatted accordingly for each mode. Examples of the versions used for Computer Assisted Telephone Interviews (CATI) and Computer Assisted Personal Interviews (CAPI) can be found in Appendix A. The internet instrument is very similar in appearance to the mail version.

Figure 1. Control (left) and Test (right) Versions of the Types of Computers Question



9 **At this house, apartment, or mobile home – do you or any member of this household own or use any of the following computers?**

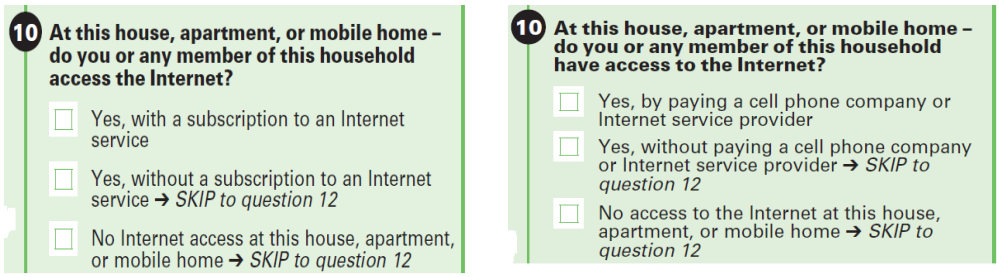
- *EXCLUDE GPS devices, digital music players, and devices with only limited computing capabilities, for example: household appliances.*

	Yes	No
a. Desktop, laptop, netbook, or notebook computer	☐	☐
b. Handheld computer, smart mobile phone, or other handheld wireless computer	☐	☐
c. Some other type of computer *Specify*	☐	☐

9 **At this house, apartment, or mobile home – do you or any member of this household own or use any of the following types of computers?**

	Yes	No
a. Desktop or laptop	☐	☐
b. Smartphone	☐	☐
c. Tablet or other portable wireless computer	☐	☐
d. Some other type of computer *Specify*	☐	☐

Figure 2. Control (left) and Test (right) Versions of the Internet Access Question



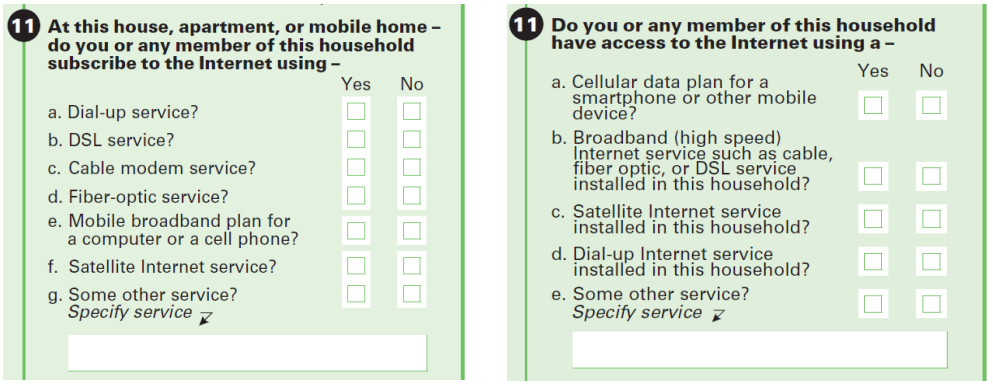
10 **At this house, apartment, or mobile home – do you or any member of this household access the Internet?**

- ☐ Yes, with a subscription to an Internet service
- ☐ Yes, without a subscription to an Internet service → *SKIP to question 12*
- ☐ No Internet access at this house, apartment, or mobile home → *SKIP to question 12*

10 **At this house, apartment, or mobile home – do you or any member of this household have access to the Internet?**

- ☐ Yes, by paying a cell phone company or Internet service provider
- ☐ Yes, without paying a cell phone company or Internet service provider → *SKIP to question 12*
- ☐ No access to the Internet at this house, apartment, or mobile home → *SKIP to question 12*

Figure 3. Control (left) and Test (right) Versions of the Internet Subscription Question



11 **At this house, apartment, or mobile home – do you or any member of this household subscribe to the Internet using –**

	Yes	No
a. Dial-up service?	☐	☐
b. DSL service?	☐	☐
c. Cable modem service?	☐	☐
d. Fiber-optic service?	☐	☐
e. Mobile broadband plan for a computer or a cell phone?	☐	☐
f. Satellite Internet service?	☐	☐
g. Some other service? *Specify service*	☐	☐

11 **Do you or any member of this household have access to the Internet using a –**

	Yes	No
a. Cellular data plan for a smartphone or other mobile device?	☐	☐
b. Broadband (high speed) Internet service such as cable, fiber optic, or DSL service installed in this household?	☐	☐
c. Satellite Internet service installed in this household?	☐	☐
d. Dial-up Internet service installed in this household?	☐	☐
e. Some other service? *Specify service*	☐	☐

1.4 Research Questions

The following research questions were formulated to guide the analyses of the questions involving Computer and Internet Use. The analyses assess how the test version of the questions perform compared to the control version in the following ways: how often the respondents answered the question, how the responses affect the resulting estimates, and the consistency and accuracy of the responses.[4]

1.4.1 Item Missing Data Rates

1. Is the item missing data rate for the types of computers question as a whole lower for the test treatment than for the control treatment?
2. Is the item missing data rate for each individual computer type lower for the test treatment than for the control treatment?
3. Is the item missing data rate for the internet access question lower for the test treatment than for the control treatment?
4. In the mail mode, is the proportion of households with multiple responses to the internet access question different between the test and control treatments?
5. Is the item missing data rate for the internet subscription type question as a whole lower for the test treatment than for the control treatment?
6. Is the item missing data rate for each individual subscription type lower for the test treatment than for the control treatment?

1.4.2 Response Proportions

7. Is the proportion of "Yes" responses for the first computer category (*Desktop/Laptop*) in the test treatment the same as the control treatment proportion?
8. Is the combined proportion of "Yes" responses for the second and third computer categories in test treatment (*Smartphone/Tablet*) greater than the proportion of "Yes" responses for the control treatment second category (*Handheld computer*)?
9. Do the changes to the types of computers question decrease the proportion in the "*Some other*" type of computer category?
10. Is the estimated proportion of households with internet access with a subscription higher in the test treatment than in the control treatment?
11. Is the estimated proportion of households without a subscription ("Access without an internet subscription" combined with "No internet access") lower in the test treatment than in the control treatment?
12. Among households that reported having a handheld device ("Smartphone" plus "Tablet" categories in test) on the types of computers question, is the proportion of those who also reported having access with a paid internet subscription higher in the test treatment than in the control treatment?
13. Is the proportion of "Dial-up" internet service the same for both treatments?

[4] Although not part of the key research, comparisons were also made to benchmark estimates from the Current Population Survey (CPS) Computer and Internet Use Supplement and surveys conducted by the Pew Research Center. Research questions, methodology, and results on benchmarks can be found in Appendix C.

14. Is the proportion of "Yes" responses obtained by collapsing the control categories of "DSL," "Cable," and "Fiber-optic" the same as the proportion of "Yes" responses for the test treatment category of "Broadband (high speed)?"
15. Is the proportion of "Cellular data" higher in the test treatment than "Mobile broadband plan" is in control?
16. Is the proportion of "Satellite" internet services the same for both treatments?
17. Is the proportion of "Some other service" in the test treatment less than or equal to the proportion in the control treatment?
18. Among households that reported having a smartphone or tablet computer in the types of computers question, is the proportion reporting "Yes" to "Mobile broadband" higher in test than in control?

1.4.3 Response Error

19. Are the measures of response reliability (gross difference rate and index of inconsistency) for each computer type category better for the test treatment than for the control treatment?
20. Are the measures of response reliability (gross difference rate and index of inconsistency) better for the test treatment than for the control treatment for the internet access question?
21. Are the measures of response reliability (gross difference rate and index of inconsistency) for each internet subscription type better for the test treatment than for the control treatment?

2 METHODOLOGY

2.1 Sample Design

The 2016 ACS Content Test consisted of a nationally representative sample of 70,000 residential addresses in the United States, independent of the production ACS sample. The Content Test sample universe did not include GQs, nor did it include housing units in Alaska, Hawaii, or Puerto Rico.[5] The sample design for the Content Test was largely based on the ACS production sample design with some modifications to better meet the test objectives.[6] The modifications included adding an additional level of stratification by stratifying addresses into high and low self-response areas, oversampling addresses from low self-response areas to ensure equal response from both strata, and sampling units as pairs.[7] The high and low self-response strata were defined based on ACS self-response rates at the tract level. Sampled pairs were formed by first systematically sampling an address within the defined sampling stratum and then pairing that address with the address listed next in the geographically sorted list. Note that the pair was likely not neighboring addresses. One member of the pair was randomly assigned to receive the

[5] Alaska and Hawaii were excluded for cost reasons. GQs and Puerto Rico were excluded because the sample sizes required to produce reliable estimates would be overly large and burdensome, as well as costly.

[6] The ACS production sample design is described in Chapter 4 of the ACS Design and Methodology report (U.S. Census Bureau, 2014).

[7] Tracts with the highest response rate based on data from the 2013 and 2014 ACS were assigned to the high response stratum in such a way that 75 percent of the housing units in the population (based on 2010 Census estimates) were in the high response areas; all other tracts were designated in the low response strata. Self-response rates were used as a proxy for overall cooperation. Oversampling in low response areas helps to mitigate larger variances due to CAPI subsampling. This stratification at the tract level was successfully used in previous ACS Content Tests, as well as the ACS Voluntary Test in 2003.

control version of the question and the other member was assigned to receive the test version of the question, thus resulting in a sample of 35,000 control cases and 35,000 test cases.

As in the production ACS, if efforts to obtain a response by mail or telephone were unsuccessful, attempts were made to interview in person a sample of the remaining nonresponding addresses (see Section 2.2 Data Collection for more details). Addresses were sampled at a rate of 1-in-3, with some exceptions that were sampled at a higher rate.[8] For the Content Test, the development of workload estimates for CATI and CAPI did not take into account the oversampling of low response areas. This oversampling resulted in a higher than expected workload for CATI and CAPI and therefore required more budget than was allocated. To address this issue, the CAPI sampling rate for the Content Test was adjusted to meet the budget constraint.

2.2 Data Collection

The field test occurred in parallel with the data collection activities for the March 2016 ACS production panel, using the same basic data collection protocol as production ACS with a few differences as noted below. The data collection protocol consisted of three main data collection operations: 1) a six-week mailout period, during which the majority of internet and mailback responses were received; 2) a one-month CATI period for nonresponse follow-up; and 3) a one-month CAPI period for a sample of the remaining nonresponse. Internet and mailback responses were accepted until three days after the end of the CAPI month.

As indicated earlier, housing units included in the Content Test sample were randomly assigned to a control or test version of the questions. CATI interviewers were not assigned specific cases; rather, they worked the next available case to be called and therefore conducted interviews for both control and test cases. CAPI interviewers were assigned Content Test cases based on their geographic proximity to the cases and therefore could also conduct both control and test cases.

The ACS Content Test's data collection protocol differed from the production ACS in a few significant ways. The Content Test analysis did not include data collected via the Telephone Questionnaire Assistance (TQA) program since those who responded via TQA used the ACS production TQA instrument. The Content Test excluded the telephone Failed Edit Follow-Up (FEFU) operation.[9] Furthermore, the Content Test had an additional telephone reinterview operation used to measure response reliability. We refer to this telephone reinterview component as the Content Follow-Up, or CFU. The CFU is described in more detail in Section 2.3.

ACS production provides Spanish-language versions of the internet, CATI, and CAPI instruments, and callers to the TQA number can request to respond in Spanish, Russian, Vietnamese, Korean, or Chinese. The Content Test had Spanish-language automated instruments; however, there were no paper versions of the Content Test questionnaires in

[8] The ACS production sample design for CAPI follow-up is described in Chapter 4, Section 4.4 of the ACS Design and Methodology report (U.S. Census Bureau, 2014).

[9] In ACS production, paper questionnaires with an indication that there are more than five people in the household or questions about the number of people in the household, and self-response returns that are identified as being vacant or a business or lacking minimal data are included in FEFU. FEFU interviewers call these households to obtain any information the respondent did not provide.

Spanish.[10] Any case in the Content Test sample that completed a Spanish-language internet, CATI, or CAPI response was included in analysis. However, if a case sampled for the Content Test called TQA to complete an interview in Spanish or any other language, the production interview was conducted and the response was excluded from the Content Test analysis. This was due to the low volume of non-English language cases and the operational complexity of translating and implementing several language instruments for the Content Test. CFU interviews for the Content Test were conducted in either Spanish or English. The practical need to limit the language response options for Content Test respondents is a limitation to the research, as some respondents self-selected out of the test.

2.3 Content Follow-Up

For housing units that completed the original interview, a CFU telephone reinterview was also conducted to measure response error.[11] A comparison of the original interview responses and the CFU reinterview responses was used to answer research questions about response error and response reliability.

A CFU reinterview was attempted with every household that completed an original interview for which there was a telephone number. A reinterview was conducted no sooner than two weeks (14 calendar days) after the original interview. Once the case was sent to CFU, it was to be completed within three weeks. This timing balanced two competing interests: (1) conducting the reinterview as soon as possible after the original interview to minimize changes in truth between the two interviews, and (2) not making the two interviews so close together that the respondents were simply recalling their previous answers. Interviewers made two call attempts to interview the household member who originally responded, but if that was not possible, the CFU reinterview was conducted with any other eligible household member (15 years or older).

The CFU asked basic demographic questions and a subset of housing and detailed person questions that included all of the topics being tested, with the exception of Telephone Service, and any questions necessary for context and interview flow to set up the questions being tested.[12] All CFU questions were asked in the reinterview, regardless of whether or not a particular question was answered in the original interview. Because the CFU interview was conducted via telephone, the wording of the questions in CFU followed the same format as the CATI nonresponse interviews. Housing units assigned to the control version of the questions in the original interview were asked the control version of the questions in CFU; housing units assigned to the test version of the questions in the original interview were asked the test version of the questions in CFU. The only exception was for retirement, survivor, and disability income, for which a different set of questions was asked in CFU.[13]

[10] In the 2014 ACS, respondents requested 1,238 Spanish paper questionnaires, of which 769 were mailed back. From that information, we projected that fewer than 25 Spanish questionnaires would be requested in the Content Test.

[11] Throughout this report the "original interview" refers to responses completed via paper questionnaire, internet, CATI, or CAPI.

[12] Because the CFU interview was conducted via telephone, the Telephone Service question was not asked. We assume that CFU respondents have telephone service.

[13] Refer to the 2016 ACS Content Test report on Retirement Income for a discussion on CFU questions for survivor, disability, and retirement income.

2.4 Analysis Metrics

This section describes the metrics used to assess the revised versions of the computer and internet use question. The metrics include the item missing data rate, response distributions, response error, and other metrics. This section also describes the methodology used to calculate unit response rates and standard errors for the test.

All Content Test data were analyzed without imputation due to our interest in how question changes or differences between versions of new questions affected "raw" responses, not the final edited variables. Some editing of responses was done for analysis purposes, such as collapsing response categories or modes together or calculating a person's age based on his or her date of birth.

All estimates from the ACS Content Test were weighted. Analysis involving data from the original interviews used the final weights that take into account the initial probability of selection (the base weight) and CAPI subsampling. For analysis involving data from the CFU interviews, the final weights were adjusted for CFU nonresponse to create CFU final weights.

The significance level for all hypothesis tests is $\alpha = 0.1$. Since we are conducting numerous comparisons between the control and test treatments, there is a concern about incorrectly rejecting a hypothesis that is actually true (a "false positive" or Type I error). The overall Type I error rate is called the familywise error rate and is the probability of making one or more Type I errors among all hypotheses tested simultaneously. When adjusting for multiple comparisons, the Holm-Bonferroni method was used (Holm, 1979).

2.4.1 Unit Response Rates and Demographic Profile of Responding Households

The unit response rate is generally defined as the proportion of sample addresses eligible to respond that provided a complete or sufficient partial response.[14] Unit response rates from the original interview are an important measure to look at when considering the analyses in this report that compare responses between the control and test versions of the survey questionnaire. High unit response rates are important in mitigating potential nonresponse bias.

For both control and test treatments, we calculated the overall unit response rate (all modes of data collection combined) and unit response rates by mode: internet, mail, CATI, and CAPI. We also calculated the total self-response rate by combining internet and mail modes together. Some Content Test analyses focused on the different data collection modes for topic-specific evaluations, thus we felt it was important to include each mode in the response rates section. In addition to those rates, we calculated the response rates for high and low response areas because analysis for some Content Test topics was done by high and low response areas. Using the Census Bureau's Planning Database (U.S. Census Bureau, 2016), we defined these areas at the tract level based on the low response score.

[14] A response is deemed a "sufficient partial" when the respondent gets to the first question in the detailed person questions section for the first person in the household.

The universe for the overall unit response rates consists of all addresses in the initial sample (70,000 addresses) that were eligible to respond to the survey. Some examples of addresses ineligible for the survey were a demolished home, a home under construction, a house or trailer that was relocated, or an address determined to be a permanent business or storage facility. The universe for self-response (internet and mail) rates consists of all mailable addresses that were eligible to respond to the survey. The universe for the CATI response rate consists of all nonrespondents at the end of the mailout month from the initial survey sample that were eligible to respond to the survey and for whom we possessed a telephone number. The universe for the CAPI response rates consists of a subsample of all remaining nonrespondents (after CATI) from the initial sample that were eligible to respond to the survey. Any nonresponding addresses that were sampled out of CAPI were not included in any of the response rate calculations.

We also calculated the CFU interview unit response rate overall and by mode of data collection of the original interview and compared the control and test treatments because response error analysis (discussed in Section 2.4.4) relies upon CFU interview data. Statistical differences between CFU response rates for control and test treatments will not be taken as evidence that one version is better than the other. For the CFU response rates, the universe for each mode consists of housing units that responded to the original questionnaire in the given mode (internet, mail, CATI, or CAPI) and were eligible for the CFU interview. We expected the response rates to be similar between treatments; however, we calculated the rates to verify that assumption.

Another important measure to look at in comparing experimental treatments is the demographic profile of the responding households in each treatment. The Content Test sample was designed with the intention of having respondents in both control and test treatments exhibit similar distributions of socioeconomic and demographic characteristics. Similar distributions allow us to compare the treatments and conclude that any differences are due to the experimental treatment instead of underlying demographic differences. Thus, we analyzed distributions for data from the following response categories: *age, sex, educational attainment,* and *tenure.* The topics of *race, Hispanic origin,* and *relationship* are also typically used for demographic analysis; however, those questions were modified as part of the Content Test, so we could not include them in the demographic profile. Additionally, we calculated *average household size* and the *language of response* for the original interview.[15]

For response distributions, we used Rao-Scott chi-square tests of independence to determine statistical differences between control and test treatments (Rao & Scott, 1987). If the distributions were significantly different, we performed additional testing on the differences for each response category. To control for the overall Type I error rate for a set of hypotheses tested simultaneously, we performed multiple-comparison procedures with the Holm-Bonferroni method (Holm, 1979). A family for our response distribution analysis was the set of p-values for the overall characteristic categories (*age, sex, educational attainment,* and *tenure*) and the set of p-values for a characteristic's response categories if the response distributions were found to have statistically significant differences. To determine statistical differences for *average household size* and the *language of response* of the original interview we performed two-tailed hypothesis tests.

[15] Language of response analysis excludes paper questionnaire returns because there was only an English questionnaire.

For all response-related calculations mentioned in this section, addresses that were either sampled out of the CAPI data collection operation or that were deemed ineligible for the survey were not included in any of the universes for calculations. Unmailable addresses were also excluded from the self-response universe. For all unit response rate estimates, differences, and demographic response analysis, we used replicate base weights adjusted for CAPI sampling (but not adjusted for CFU nonresponse).

2.4.2 Item Missing Data Rates

Respondents leave items blank for a variety of reasons including not understanding the question (clarity), their unwillingness to answer a question as presented (sensitivity), and their lack of knowledge of the data needed to answer the question. The item missing data rate (for a given item) is the proportion of eligible units, housing units for household-level items or persons for person-level items, for which a required response (based on skip patterns) is missing.

We calculated and compared the item missing data rates between control and test for all of the Computer and Internet Use questions. All respondents were required to answer the types of computers and internet access questions. Only those units that responded that they had internet access "with a subscription to an internet service" for control or "by paying a cell phone company or internet service provider" for test were required to answer the question about types of internet subscriptions. Statistical significance of differences between versions was determined using two-tailed t-tests.

Types of Computers
The percentage of eligible housing units that did not provide a response in the control treatment was compared to the corresponding percentage from the test treatment. In addition to evaluating the overall question, missing data rates for the new test categories were compared individually to the control categories, resulting in three tests of item missing data rates on individual computer types. On mail and internet questionnaires, missing responses were those where no boxes were marked. In CATI and CAPI instruments, a response of either "Don't Know" or "Refused" was considered missing. Responses where "Some other type of computer" was marked but no write-in was provided were not considered missing. If one type of computer was marked "Yes," any other type of computer that was left blank was considered to be a "No" instead of a missing answer.

Internet Access
The percentage of eligible housing units that did not provide a response to this question in the control treatment was compared to the corresponding percentage from the test treatment. On mail and internet questionnaires, missing responses were those where no boxes were marked. In the CATI and CAPI instruments, a response of either "Don't Know" or "Refused" was considered missing.

A limitation of the mail questionnaire version of the internet access question is that a respondent may erroneously mark more than one box as an answer to the question. If more than one box was marked then the answer was considered missing, since we cannot assume which answer is the correct one. We also calculated the number of times a respondent checked multiple boxes for the internet access question. We compared the proportions of responses with multiple marks, using

adjusted weights, between control and test. We expected the percentage of multiple responses of the test version to be the same as or lower than the control version.

Internet Subscription
The percentage of eligible housing units that did not provide a response to this question in the control treatment was compared to the corresponding percentage from the test treatment. As with the computer question, we needed an assessment of overall nonresponse as well as nonresponse for individual components. On mail and internet questionnaires, missing responses were those where no boxes were marked. In CATI and CAPI instruments, a response of either "Don't Know" or "Refused" was considered missing. Responses where "Some other service" was marked but no write-in was provided were not considered missing. If one type of internet subscription was marked "Yes," any other type that was left blank was considered to be a "No" instead of a missing answer.

2.4.3 Response Proportions

Comparing the proportion of the response categories between the control version of a question and the test version of a question allows us to assess whether the question change affects the resulting estimates.

Proportion estimates were calculated as:

$$\text{Category proportion} = \frac{\text{weighted count of valid responses in category}}{\text{weighted count of all valid responses}}$$

Types of Computers
The control category "Desktop, laptop, netbook, or notebook computer" was compared to the test category "Desktop or laptop" using a two-tailed t-test as the percentages were not expected to differ. The control category for "Handheld computer" was compared to the combined test categories for "Smartphone" and "Tablet or other portable wireless computer." We used a one-tailed t-test because we expected the test to show a greater percentage of households with smartphone and tablet ownership, due to the updated changes to the categories. A straight comparison was made between control and test in the category of "Some other type of computer." This analysis only involved the checkbox and did not check for the presence or content of a write-in. We compared the "some other type of computer" category between treatments using a two-tailed t-test. Ideally, we expected to see a lower percentage of "Some other type of computer" responses in test than in control, as the new version added "Tablet" and isolated "Smartphone" to its own category; however, similar proportions were also acceptable.

Internet Access
Both internet access categories were compared with two-tailed t-tests. Although we expected the test treatment to have a greater percentage of respondents with internet access due to the inclusion of "paying a cell phone company" in the question, we considered an outcome of similar proportions to be acceptable. Also, it was expected that the test version would have a lower percentage of respondents reporting that they had access without a subscription or that they did not have internet access; however we considered it acceptable if the percentages were similar.

When comparing internet access among households reporting a handheld device, the control universe included all households with a handheld device from the types of computers question while the test universe included all households with either a smartphone or tablet. We compared the percentage of each universe reporting a paid internet subscription (the first box in each version of the internet access question). We compared the proportions using a two-tailed t-test.

Internet Subscription
The control categories of "Dial-up" and "Satellite" were compared to the corresponding test categories using two-tailed t-tests. Since these categories did not change in the test version, they were expected to have similar percentages of "Yes" responses. The control categories of "DSL," "Cable," and "Fiber-optic" were combined and compared to the test category of "Broadband (high speed)" using a two-tailed t-test. This percentage comparison was also expected to be about the same for control and test. The control category of "Mobile broadband plan" was compared to the test category of "Cellular data plan" using a one-tailed t-test. Due to the change in terminology, the test version was expected to result in a higher percentage of mobile broadband subscribers. The category of "Some other service" was compared between control and test using a two-tailed t-test. Because of the clarity of the new categories in the test version, we expected to receive a similar or lower percentage of respondents reporting in the "Some other service" category.

When assessing mobile broadband among households reporting a handheld device, similar to the analysis for internet access, we compared control households with a handheld device to test households with either a smartphone or tablet. We compared the proportions using a one-tailed t-test.

2.4.4 Response Error

Response error occurs for a variety of reasons, such as flaws in the survey design, misunderstanding of the questions, misreporting by respondents, or interviewer effects. There are two components of response error: response bias and simple response variance. Response bias is the degree to which respondents consistently answer a question incorrectly. Simple response variance is the degree to which respondents answer a question inconsistently. A question has good response reliability if respondents tend to answer the question consistently. Re-asking the same question of the same respondent (or housing unit) allows us to measure response variance.

We measured simple response variance by comparing valid responses to the CFU reinterview with valid responses to the corresponding original interview.[16] The Census Bureau has frequently used content reinterview surveys to measure simple response variance for large demographic data collection efforts, including the 2010 ACS Content Test, and the 1990, 2000, and 2010 decennial censuses (Dusch & Meier, 2012).

[16] A majority of the CFU interviews were conducted with the same respondent as the original interview (see the Limitations section for more information).

The following measures were used to evaluate consistency:

- Gross difference rate (GDR)
- Index of inconsistency (IOI)
- L-fold index of inconsistency (IOI_L)

The first two measures – GDR and IOI – were calculated for individual response categories. The L-fold index of inconsistency was calculated for questions that had three or more mutually exclusive response categories, as a measure of overall reliability for the question.

The GDR, and subsequently the simple response variance, are calculated using the following table and formula.

Table 1. Interview and Reinterview Counts for Each Response Category Used for Calculating the Gross Difference Rate and Index of Inconsistency

	Original Interview "Yes"	Original Interview "No"	**Reinterview Totals**
CFU Reinterview "Yes"	A	b	**a + b**
CFU Reinterview "No"	C	d	**c + d**
Original Interview Totals	**a + c**	**b + d**	**n**

Where a, b, c, d, and n are defined as follows:

a = weighted count of units in the category of interest for both the original interview and reinterview
b = weighted count of units NOT in the category of interest for the original interview, but in the category for the reinterview
c = weighted count of units in the category of interest for the original interview, but NOT in the category for the reinterview
d = weighted count of units NOT in the category of interest for either the original interview or the reinterview
n = total units in the universe = a + b + c + d.

The GDR for a specific response category is the percent of inconsistent answers between the original interview and the reinterview (CFU). We calculate the GDR for a response category as

$$\text{GDR} = \frac{(b + c)}{n} \times 100$$

Statistical significance between the GDR for a specific response category between the control and test treatments is determined using a two-tailed t-test.

In order to define the IOI, we must first discuss the variance of a category proportion estimate. If we are interested in the true proportion of a total population that is in a certain category, we can use the proportion of a survey sample in that category as an estimate. Under certain reasonable assumptions, it can be shown that the total variance of this proportion estimate is the sum of two

components, sampling variance (SV) and simple response variance (SRV). It can also be shown that an unbiased estimate of SRV is half of the GDR for the category (Flanagan, 1996).

SV is the part of total variance resulting from the differences among all the possible samples of size *n* one might have selected. SRV is the part of total variance resulting from the aggregation of response error across all sample units. If the responses for all sample units were perfectly consistent, then SRV would be zero, and the total variance would be due entirely to SV. As the name suggests, the IOI is a measure of how much of the total variance is due to inconsistency in responses, as measured by SRV and is calculated as:

$$\mathrm{IOI} = \frac{\mathrm{n(b + c)}}{\mathrm{(a + c)(c + d) + (a + b)(b + d)}} \times 100$$

Per the Census Bureau's general rule, index values of less than 20 percent indicate low inconsistency, 20 to 50 percent indicate moderate inconsistency, and over 50 percent indicate high inconsistency.

An IOI is computed for each response category and an overall index of inconsistency, called the L-fold index of inconsistency, is reported for the entire distribution. The L-fold index is a weighted average of the individual indexes computed for each response category.

When the sample size is small, the reliability estimates are unstable. Therefore, we do not report the IOI and GDR values for categories with a small sample size, as determined by the following formulas: $2a + b + c < 40$ or $2d + b + c < 40$, where a, b, c, and d are unweighted counts as shown in Table 1 above (see Flanagan 1996, p. 15).

The measures of response error assume that those characteristics in question did not change between the original interview and the CFU interview. To the extent that this assumption is incorrect, we assume that it is incorrect at similar rates between the control and test treatments. An example of this could be a question on ownership of computer devices. For instance, a household that did not report having a tablet originally might have acquired one before the CFU interview and then accurately reported a different response than the original.

In calculating the IOI reliability measures, the assumption is that the expected value of the error in the original interview is the same as in the CFU reinterview. This assumption of parallel measures is necessary for the SRV and IOI to be valid. In calculating the IOI measures for this report, we found this assumption was not met for the response categories specified in the limitations section (see Section 4).

Biemer (2011, pp. 56-58) provides an example where the assumption of parallel measures is not met, but does not provide definitive guidelines for addressing it. In Biemer's concluding remarks, he states, "...both estimates of reliability are biased to some extent because of the failure of the parallel assumptions to hold."

Flanagan (2001) addresses this bias problem and offers the following adjustment to the IOI formula:

$$IOI_{testimate} = \frac{\frac{n^2(b+c) - n(c-b)^2}{n-1}}{(a+c)(c+d) + (a+b)(b+d)} \times 100$$

This formula was tested on selected topics in the 2016 ACS Content Test. The $IOI_{testimate}$ resulted in negligible reduction in the IOI values. For this reason, we did not recalculate the IOI values using $IOI_{testimate}$. Similar to Biemer (2011, p. 58), we acknowledge that for some cases, the estimate of reliability is biased to some extent.

For the Computer and Internet Use content, analysis examined the reliability—GDRs and IOIs—of each category of the types of computers, internet access, and internet subscription questions. When analyzing the types of computers question, categories for "Smartphone" and "Tablet" in the test version were combined for comparison with the "Handheld" category in the control version. For the internet subscription item, categories for "DSL", "Cable", and "Fiber-optic" in the control version were aggregated for comparison with the "Broadband (high speed)" category in the test version. The specific content of the write-in fields for "Some other computer" and "Some other service" were not assessed for reliability.

In addition, the IOI_L for the internet access item was determined to estimate overall reliability for the question as a whole. It is not appropriate to calculate the IOI_L for the types of computers or internet subscription questions, as the categories for these items are not mutually exclusive. For all Computer and Internet Use items, statistical significance between the GDRs and IOIs of each version were determined using two-tailed t-tests.

2.4.5 Standard Error Calculations

We estimated the variances of the estimates using the Successive Differences Replication (SDR) method with replicate weights, the standard method used in the ACS (see U.S. Census Bureau, 2014, Chapter 12). We calculated the variance for each rate and difference using the formula below. The standard error of the estimate (X_0) is the square root of the variance:

$$Var(X_0) = \frac{4}{80}\sum_{r=1}^{80}(X_r - X_0)^2$$

where:

X_0 = the estimate calculated using the full sample,
X_r = the estimate calculated for replicate r.

3 KEY RESEARCH CRITERIA FOR COMPUTER AND INTERNET USE

Before fielding the 2016 ACS Content Test, we identified which of the metrics would be given higher importance in determining which version of the question yielded the best quality of data for each topic. The following tables identify the research questions and associated metrics and criteria in priority order.

Table 2. Key Research Criteria for Types of Computers Question

Research Questions	Research Criteria In Order of Priority
19	The reliability for the test version should be the same as or greater than the control version, especially for smartphone and tablet users as compared to handheld device users.
8	The proportion of responses indicating smartphone or tablet use should be greater in the test treatment than the proportion of responses from control that indicate handheld device use.
1, 2	The item missing data rates for the test treatment should be lower than or the same as the control treatment.
9	The proportion of responses indicating use of some other type of computer for the test treatment should be the same as or lower than the control treatment.
7	Additionally, the proportion of responses indicating desktop or laptop use should be the same between control and test treatments.

Table 3. Key Research Criteria for Internet Access Question

Research Questions	Research Criteria In Order of Priority
12	Among households with a smartphone or tablet (handheld in control), the proportion having an internet subscription for the test treatment should be the same as or higher than the control treatment.
20	The reliability for the test version should be the same as or greater than the control version.
3	The item missing data rates for the test version should be lower than or the same as the control version.
10, 11	Among all households, the proportion having an internet subscription for the test treatment should be the same as or higher than the control treatment. Similarly, the proportion without a subscription in the test should be the same as or lower than the control proportion.
4	In the mail mode, the proportion of households with multiple responses to the internet Access question in the test should be the same as or lower than the control proportion.

Table 4. Key Research Criteria for Internet Subscription Question

Research Questions	Research Criteria In Order of Priority
18	Among households with a smartphone or tablet (handheld in control), the proportion having a cellular data subscription in the test treatment should be higher than the proportion with mobile broadband in the control treatment.
21	The reliability for the test version should be the same as or greater than the control version, when aggregating categories appropriately.
13, 16, 17, 15	The proportion of "Dial-up" and "Satellite" responses should be the same in the test as in the control. Similarly, the proportion of "Some other service" responses for the test version should be the same as or lower than the control version. Finally, the proportion of "Cellular data" responses in the test should be higher than "Mobile broadband" responses in the control.
5, 6	The item missing data rates for the test treatment should be the same as or lower than the control treatment, when measured as the failure to mark any element of the question. Similarly, the item missing data rates in the test for each individual subscription type should be the same as or lower than the control rates.
14	The proportion of "Yes" responses obtained by collapsing the control categories "DSL," "Cable," and "Fiber-optic" should be the same as the proportion of "Yes" responses for the test treatment category of "Broadband (high speed)."

4 LIMITATIONS

CATI and CAPI interviewers were assigned control and test treatment cases, as well as production cases. The potential risk of this approach is the introduction of a cross-contamination or carry-over effect due to the same interviewer administering multiple versions of the same question item. Interviewers are trained to read the questions verbatim to minimize this risk, but there still exists the possibility that an interviewer may deviate from the scripted wording of one question version to another. This could potentially mask a treatment effect from the data collected.

Interviews were only conducted in English and Spanish. Respondents who needed language assistance in another language were not able to participate in the test. Additionally, the 2016 ACS Content Test was not conducted in Alaska, Hawaii, or Puerto Rico. Any conclusions drawn from this test may not apply to these areas or populations.

For statistical analysis specific to the mail mode, there may be bias in the results because of unexplained unit response rate differences between the control and test treatments.

We were not able to conduct demographic analysis by relationship status, race, or ethnicity because these topics were tested as part of the Content Test.

The CFU reinterview was not conducted in the same mode of data collection for households that responded by internet, by mail, or by CAPI in the original interview since CFU interviews were only administered using a CATI mode of data collection. As a result, the data quality measures derived from the reinterview may include some bias due to the differences in mode of data collection.

To be eligible for a CFU reinterview, respondents needed to either provide a telephone number in the original interview or have a telephone number available to the Census Bureau through reverse address look up. As a result, 2,284 of the responding households (11.8 percent with a standard error of 0.2) from the original control interviews and 2,402 of the responding households (12.4 percent with a standard error of 0.2) from the original test interviews were not eligible for the CFU reinterview. The difference between the control and test treatments is statistically significant (p-value=0.06).

Although we reinterviewed the same person who responded in the original interview when possible, we interviewed a different member of the household in the CFU for 7.5 percent (standard error of 0.4) of the CFU cases for the control treatment and 8.4 percent (standard error of 0.5) of the CFU cases for the test treatment.[17] The difference between the test and control treatments is not statistically significant (p-value=0.26). This means that differences in results between the original interview and the CFU for these cases could be due in part to having different people answering the questions. However, those changes were not statistically significant between the control and test treatments and should not impact the conclusions drawn from the reinterview.

The 2016 ACS Content Test does not include the production weighting adjustments for seasonal variations in ACS response patterns, nonresponse bias, and under-coverage bias. As a result, any estimates derived from the Content Test data do not provide the same level of inference as the production ACS and cannot be compared to production estimates.

In developing initial workload estimates for CATI and CAPI, we did not take into account the fact that we oversampled low response areas as part of the Content Test sample design. Therefore, workload and budget estimates were too low. In order to stay within budget, the CAPI workload was subsampled more than originally planned. This caused an increase in the variances for the analysis metrics used.

An error in addressing and assembling the materials for the 2016 ACS Content Test caused some Content Test cases to be mailed production ACS questionnaires instead of Content Test questionnaires. There were 49 of these cases that returned completed questionnaires, and they were all from the test treatment. These cases were excluded from the analysis. Given the small number of cases affected by this error, there is very little effect on the results.

Questionnaire returns were expected to be processed and keyed within two weeks of receipt. Unfortunately, a check-in and keying backlog prevented this requirement from being met, thereby delaying eligible cases from being sent to CFU on a schedule similar to the other modes. Additionally, the control treatment questionnaires were processed more quickly in keying than the test treatment questionnaires resulting in a longer delay for test mail cases to be eligible for CFU. On average, it took 18 days for control cases to become eligible for CFU; it took 20 days for test cases. The difference is statistically significant. This has the potential to impact the response reliability results.

[17] This is based on comparing the first name of the respondent between the original interview and the CFU interview. Due to a data issue, we were not able to use the full name to compare.

The assumption of parallel measures for the GDR and IOI calculations was not met for the following categories: some other type of computer, access with a subscription, access without a subscription, and mobile broadband internet service. For these categories, the GDR and IOI estimates are biased to some extent.

5 RESEARCH QUESTIONS AND RESULTS

This section presents the results from the analyses of the 2016 ACS Content Test data for the Computer and Internet Use questions. An analysis of unit response rates is presented first followed by topic-specific analyses. For the topic-specific analyses, each research question is restated, followed by corresponding data and a brief summary of the results.

5.1 Unit Response Rates and Demographic Profile of Responding Households

This section provides results for unit response rates for both control and test treatments for the original Content Test interview and for the CFU interview. It also provides results of a comparison of socioeconomic and demographic characteristics of respondents in both control and test treatments.

5.1.1 Unit Response Rates for the Original Content Test Interview

The unit response rate is generally defined as the proportion of sample addresses eligible to respond that provided a complete or sufficient partial response. We did not expect the unit response rates to differ between treatments. This is important because the number of unit responses should also affect the number of item responses we receive for analyses done on specific questions on the survey. Similar item response universe sizes allow us to compare the treatments and conclude that any differences are due to the experimental treatment instead of differences in the populations sampled for each treatment.

Table 5 shows the unit response rates for the original interview for each mode of data collection (internet, mail, CATI, and CAPI), all modes combined, and both self-response modes (internet and mail combined) for the control and test treatments. When looking at the overall unit response rate (all modes combined), the difference between control (93.5 percent) and test (93.5 percent) is less than 0.1 percentage points and is not statistically significant.

Table 5. Original Interview Unit Response Rates for Control and Test Treatments, Overall and by Mode

Mode	Test Interviews	Test Percent	Control Interviews	Control Percent	Test minus Control	P-Value
All Modes	19,400	93.5 (0.3)	19,455	93.5 (0.3)	<0.1 (0.4)	0.98
Self-Response	13,131	52.9 (0.5)	13,284	53.7 (0.5)	-0.8 (0.6)	0.23
Internet	8,168	34.4 (0.4)	8,112	34.1 (0.4)	0.4 (0.6)	0.49
Mail	4,963	18.4 (0.3)	5,172	19.6 (0.3)	-1.2 (0.5)	0.01*
CATI	872	8.7 (0.4)	880	9.2 (0.4)	-0.4 (0.6)	0.44
CAPI	5,397	83.5 (0.7)	5,291	83.6 (0.6)	<0.1 (0.9)	0.96

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level. The weighted response rates account for initial sample design as well as CAPI subsampling.

When analyzing the unit response rates by mode of data collection, the only modal comparison that shows a statistically significant difference is the mail response rate. The control treatment had a higher mail response (19.6 percent) than the test treatment (18.4 percent) by 1.2 percentage points. As a result of this difference, we looked at how mail responses differed in the high and low response areas. Table 6 shows the mail response rates for both treatments in high and low response areas.[18] The difference in mail response rates appears to be driven by the difference of rates in the high response areas.

It is possible that the difference in the mail response rates between control and test is related to the content changes made to the test questions. There are some test questions that could be perceived as being too sensitive by some respondents (such as the test question relating to same-sex relationships) and some test questions that could be perceived to be too burdensome by some respondents (such as the new race questions with added race categories). In the automated modes (internet, CATI, and CAPI) there is a higher likelihood of obtaining a sufficient partial response (obtaining enough information to be deemed a response for calculations before the respondent stops answering questions) than in the mail mode. If a respondent is offended by the questionnaire or feels that the questions are too burdensome they may just throw the questionnaire away, and not respond by mail. This could be a possible explanation for the unit response rate being lower for test than control in the mail mode.

We note that differences between overall and total self-response response rates were not statistically significant. As most analysis was conducted at this level, we are confident the response rates were sufficient to conduct topic-specific comparisons between the control and test treatments and that there are no underlying response rate concerns that would impact those findings.

[18] Table B-1 (including all modes) can be found in Appendix B.

Table 6. Mail Response Rates by Designated High (HRA) and Low (LRA) Response Areas

	Test Interviews	Test Percent	Control Interviews	Control Percent	Test minus Control	P-Value
HRA	2,082	20.0 (0.4)	2,224	21.5 (0.4)	-1.5 (0.6)	0.02*
LRA	2,881	13.8 (0.3)	2,948	14.1 (0.3)	-0.3 (0.4)	0.43
Difference		6.2 (0.5)		7.4 (0.4)	-1.1 (0.7)	0.11

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level. The weighted response rates account for initial sample design as well as CAPI subsampling.

5.1.2 Unit Response Rates for the Content Follow-Up Interview

Table 7 shows the unit response rates for the CFU interview by mode of data collection of the original interview and for all modes combined, for control and test treatments. Overall, the differences in CFU response rates between the treatments are not statistically significant. The rate at which CAPI respondents from the original interview responded to the CFU interview is lower for test (34.8 percent) than for control (37.7 percent) by 2.9 percentage points. While the protocols for conducting CAPI and CFU were the same between the test and control treatments, we could not account for personal interactions that occur in these modes between the respondent and interviewer. This can influence response rates. We do not believe that the difference suggests any underlying CFU response issues that would negatively affect topic-specific response reliability analysis for comparing the two treatments.

Table 7. Content Follow-Up Interview Unit Response Rates for Control and Test Treatments, Overall and by Mode of Original Interview

Original Interview Mode	Test Interviews	Test Percent	Control Interviews	Control Percent	Test minus Control	P-Value
All Modes	7,867	44.8 (0.5)	7,903	45.7 (0.6)	-0.8 (0.8)	0.30
Internet	4,078	51.9 (0.6)	4,045	52.5 (0.7)	-0.6 (0.8)	0.49
Mail	2,202	46.4 (0.9)	2,197	44.2 (0.9)	2.1 (1.3)	0.11
CATI	369	48.9 (1.9)	399	51.5 (2.5)	-2.5 (2.9)	0.39
CAPI	1,218	34.8 (1.2)	1,262	37.7 (1.1)	-2.9 (1.6)	0.07*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level.

5.1.3 Demographic and Socioeconomic Profile of Responding Households

One of the underlying assumptions of our analyses in this report is that the sample for the Content Test was selected in such a way that responses from both treatments would be comparable. We did not expect the demographics of the responding households for control and test treatments to differ. To test this assumption, we calculated distributions for respondent data for the following response categories: *age*, *sex*, *educational attainment,* and *tenure.*[19] The

[19] We were not able to conduct demographic analysis by relationship status, race, or ethnicity because these topics were tested as part of the Content Test.

response distribution calculations can be found in Table 8. Items with missing data were not included in the calculations. After adjusting for multiple comparisons, none of the differences in the categorical response distributions shown below is statistically significant.

Table 8. Response Distributions: Test versus Control Treatment

Item	Test Percent	Control Percent	Adjusted P-Value
AGE	(n=43,236)	(n=43,325)	0.34
Under 5 years old	5.7 (0.2)	6.1 (0.2)	
5 to 17 years old	17.8 (0.3)	17.6 (0.3)	
18 to 24 years old	8.6 (0.3)	8.1 (0.3)	
25 to 44 years old	25.1 (0.3)	26.2 (0.3)	
45 to 64 years old	26.8 (0.4)	26.6 (0.4)	
65 years old or older	16.0 (0.3)	15.4 (0.3)	
SEX	(n=43,374)	(n=43,456)	1.00
Male	48.8 (0.3)	49.1 (0.3)	
Female	51.2 (0.3)	50.9 (0.3)	
EDUCATIONAL ATTAINMENT[#]	(n=27,482)	(n=27,801)	1.00
No schooling completed	1.3 (0.1)	1.2 (0.1)	
Nursery to 11th grade	8.1 (0.3)	8.0 (0.3)	
12th grade (no diploma)	1.7 (0.1)	1.6 (0.1)	
High school diploma	21.7 (0.4)	22.3 (0.4)	
GED† or alternative credential	3.5 (0.2)	3.6 (0.2)	
Some college	21.0 (0.4)	20.2 (0.4)	
Associate's degree	8.8 (0.3)	9.1 (0.3)	
Bachelor's degree	20.9 (0.4)	20.3 (0.4)	
Advanced degree	13.1 (0.3)	13.7 (0.3)	
TENURE	(n=17,190)	(n=17,236)	1.00
Owned with a mortgage	43.1 (0.6)	43.2 (0.5)	
Owned free and clear	21.1 (0.4)	21.2 (0.4)	
Rented	33.8 (0.6)	34.0 (0.5)	
Occupied without payment of rent	1.9 (0.2)	1.7 (0.1)	

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
[#]For ages 25 and older
†General Educational Development
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance testing done at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

We also analyzed two other demographic characteristics shown by the responses from the survey: *average household size* and *language of response.* The results for the remaining demographic analyses can be found in Table 9 and Table 10.

Table 9. Comparison of Average Household Size

	Test (n=17,608)	Control (n=17,694)	Test minus Control	P-value
Average Household Size (Number of People)	2.51 (<0.1)	2.52 (<0.1)	>-0.01 (<0.1)	0.76

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Significance was tested based on a two-tailed t-test at the α=0.1 level.

Table 10. Comparison of Language of Response

Language of Response	Test Percent (n=17,608)	Control Percent (n=17,694)	Test minus Control	P-value
English	96.1 (0.2)	96.2 (0.2)	<0.1 (0.3)	0.52
Spanish	2.7 (0.2)	2.6 (0.2)	<0.1 (0.2)	0.39
Undetermined	1.2 (0.1)	1.2 (0.1)	<0.1 (0.2)	0.62

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Significance was tested based on a two-tailed t-test at the α=0.1 level.

The Content Test was available in two languages, English and Spanish, for all modes except the mail mode. However, the language of response variable was missing for some responses, so we created a category called *undetermined* to account for those cases.

There are no detectable differences between control and test for *average household size* or *language of response*. There are also no detectable differences for any of the response distributions that we calculated. As a result of these analyses, it appears that respondents in both treatments do exhibit comparable demographic characteristics since none of the resulting findings is significant, which verifies our assumption of demographic similarity between treatments.

5.2 Item Missing Data Rates

Is the item missing data rate for the types of computers question as a whole lower for the test treatment than for the control treatment?

The first row of Table 11 shows the item missing data rates for the types of computers question as a whole. There are no significant differences in the items missing data rates between treatments for any of the computer type categories. This suggests that the changes made to the question do not affect item nonresponse.

Table 11. Item Missing Data Rates for Control and Test Treatments, Types of Computers Question

Item	Test Percent (n=17,588)	Control Percent (n=17,688)	Test minus Control	Adjusted P-Value
Entire question	1.3 (0.1)	1.4 (0.1)	-0.1 (0.2)	1.00
Desktop or laptop	1.5 (0.1)	1.5 (0.1)	<0.1 (0.2)	1.00
Smartphone or tablet vs. Handheld	1.5 (0.1)	1.7 (0.1)	-0.2 (0.2)	0.91
Other computer	1.7 (0.1)	2.0 (0.1)	-0.3 (0.2)	0.23

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test (test≠control) at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Is the item missing data rate for each individual computer type lower for the test treatment than for the control treatment?

Item missing data rates for each computer type category are displayed above in Table 11. Similar to what was observed for the question overall, the item missing data rate for individual categories is not significantly different for test versus control, indicating that the test version does not reduce or increase item nonresponse.

Is the item missing data rate for the internet access question lower for the test treatment than for the control treatment?

Table 12 contains information on item missing data rates for the internet access question. Item missingness is significantly lower in the test treatment (2.0 percent) than in the control treatment (2.3 percent), indicating that the test version of the question performed better in terms of item missingness. Omitting the confusing term "subscription" from the test version of the question likely made it easier for some respondents to answer.

Table 12. Item Missing Data Rates for Control and Test Treatments, Internet Access Question

Item	Test Percent (n=17,588)	Control Percent (n=17,688)	Test minus Control	P-Value
Entire question	2.0 (0.2)	2.3 (0.2)	-0.3 (0.2)	0.07*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test (test≠control) at the α=0.1 level.

In the mail mode, is the proportion of households with multiple responses to the internet access question different between the test and control treatments?

The share of households providing multiple responses to the internet access question in the mail mode is found in Table 13. We include the results of multiple responses in this section, as the internet access item is considered missing for cases marking more than one box. There is no significant difference between treatments in the proportion of households with multiple responses, indicating that the changes to the question did not affect this indicator.

Table 13. Proportion of Households with Multiple Responses on Mail Questionnaire, Internet Access Question

Test Percent (n=4,859)	Control Percent (n=5,062)	Test minus Control	P-Value
0.5 (0.1)	0.5 (0.1)	0.1 (0.2)	0.75

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Significance was tested based on a two-tailed t-test at the α=0.1 level.

Is the item missing data rate for the internet subscription type question as a whole lower for the test treatment than for the control treatment?

The first row of Table 14 displays item missing data rates for the internet subscription type question as a whole. Note that whereas the universe for the types of computers and internet access questions is all eligible housing units, the universe for the internet subscription type question is all eligible housing units that have internet access with a subscription. The item missing data rate is not significantly different between the control and test treatments.

Table 14. Item Missing Data Rates for Control and Test Treatments, Internet Subscription Type Question

Item	Test Percent (n=14,033)	Control Percent (n=13,624)	Test minus Control	Adjusted P-Value
Entire question	2.3 (0.2)	1.8 (0.2)	0.5 (0.3)	0.25
Dial-up	1.0 (0.1)	1.1 (0.1)	-0.2 (0.1)	0.29
High speed vs. DSL/Cable/Fiber-optic	1.0 (0.1)	0.9 (0.1)	<0.0 (0.1)	0.78
Cellular data plan vs. Mobile broadband	0.8 (0.1)	1.2 (0.1)	-0.4 (0.1)	0.01*
Satellite	1.0 (0.1)	1.2 (0.1)	-0.3 (0.1)	0.25
Other service	1.0 (0.1)	1.3 (0.1)	-0.3 (0.1)	0.25

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test (test ≠ control) at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Is the item missing data rate for each individual subscription type lower for the test treatment than for the control treatment?

Finally, information on missingness for individual subscription types is found above in Table 14. Of the five categorical comparisons made, the only significant difference detected in the item missing data rates is the rate for the "Cellular data plan" category. Missingness is lower in the test treatment (0.8 percent) than in the control treatment (1.2 percent), suggesting that respondents understand the phrase "Cellular data plan" better than the phrase "Mobile broadband plan."

5.3 Response Proportions

For all Computer and Internet Use questions, the universe for the response proportion analysis is households with a nonmissing response to the item of concern.

Is the proportion of "Yes" responses for the first computer category (Desktop/Laptop) in the test treatment the same as the control treatment proportion?

Table 15 displays the response proportions for each category of the types of computers question. Although we expected the same share of households to report owning or using a desktop or laptop in each treatment, results indicate that this proportion is lower in the test treatment (78.6 percent) than in the control treatment (80.7 percent). A possible explanation for the observed

difference is the introduction of the "Tablet" category to the test version of the question. In the absence of this category, some control respondents owning or using tablets (but not desktops or laptops) may have marked the category for "Desktop, laptop, netbook, or notebook computer."

Table 15. Response Proportions for Control and Test Treatments, Types of Computers Question

Category	Test Percent (n=17,329)	Control Percent (n=17,387)	Test minus Control	Alternative Hypothesis	Adjusted P-Value
Desktop or laptop	78.6 (0.4)	80.7 (0.4)	-2.1 (0.6)	T≠C	<0.01*
Smartphone or tablet vs. Handheld	82.4 (0.4)	79.8 (0.4)	2.6 (0.6)	T>C	<0.01*
Other computer	4.6 (0.2)	7.9 (0.3)	-3.3 (0.4)	T≠C	<0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference between the two rates at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method. Question allows for multiple categories to be marked, so columns will not sum to 100 percent.

Is the combined proportion of "Yes" responses for the second and third computer categories in test treatment (Smartphone/Tablet) greater than the proportion of "Yes" responses for the control treatment second category (Handheld computer)?

Looking once more at Table 15 above, as expected, results reveal that a larger proportion of test households reported owning or using a smartphone or tablet (82.4 percent), compared with the share of control households reporting a handheld computer (79.8 percent). Under the old (control) question wording, some smartphone and/or tablet owners may not have recognized the category for "Handheld computer, smart mobile phone, or other handheld wireless computer" as applying to them. Specific categories for "Smartphone" and "Tablet or other portable wireless computer" found in the new (test) wording likely are better understood by those with these devices.

Do the changes to the types of computers question decrease the proportion in the "Some other" category?

The final row of Table 15 above shows the share of households reporting some other computer in the test and control treatments. As predicted, a smaller proportion of households in the test treatment indicated that they owned or used some other computer (4.6 percent), compared with the control treatment (7.9 percent). This change is also likely due to replacing the "Handheld computer, smart mobile phone, or other handheld wireless computer" category with specific options for "Smartphone" and "Tablet or other portable wireless computer." Under the old (control) wording, some smartphone and/or tablet users may have marked the "other" category, but the new (test) wording makes it easier to find the relevant descriptive category(ies).

Is the estimated proportion of households with internet access with a subscription higher in the test treatment than in the control treatment?

Table 16 contains response proportions for the internet access question. As expected, results show that the proportion of households reporting internet access with a subscription is higher in the test treatment, at 83.8 percent, than in the control treatment, at 82.3 percent. As suggested by earlier cognitive testing, respondents likely find the term "paying," used in the test version of the question, clearer than the term "subscription," used in the control version. Also important, adding the phrase "cell phone company" likely resonated with respondents who receive internet through a cell phone provider instead of or in addition to a conventional internet service provider.

Table 16. Response Proportions for Control and Test Treatments, Internet Access Question

Category	Test Percent (n=17,171)	Control Percent (n=17,188)	Test minus Control	P-Value
Access with subscription	83.8 (0.4)	82.3 (0.4)	1.5 (0.6)	0.01*
No subscription	16.2 (0.4)	17.7 (0.4)	-1.5 (0.6)	0.01*
Total	100.0	100.0		

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test (test ≠ control) at the α=0.1 level.

Is the estimated proportion of households without a subscription ("Access without an internet subscription" combined with "No internet access") lower in the test treatment than in the control treatment?

We see (in Table 16) that at the same time that reporting of internet subscriptions was higher for test households, reporting of no internet subscription was lower among test households. Only 16.2 percent of households in the test treatment indicated having internet access without a subscription or no internet access, compared with 17.7 percent of households in the control treatment.

Among households that reported having a handheld device ("Smartphone" plus "Tablet" categories in test) on the types of computers question, is the proportion of those who also reported having access with a paid internet subscription higher in the test treatment than in the control treatment?

shows the share of households reporting access with a subscription, looking specifically at households owning a device such as a smartphone or tablet. As was observed for all households overall, those with a smartphone or tablet are more likely to report a subscription when receiving the test version of the question (92.4 percent) than when seeing the control version (90.5 percent). Thus, the revised question wording better captures internet access among portable device owners as well as for the general population.

Table 17. Proportion of Households with Smartphone or Tablet Reporting Access with a Subscription

Test Percent (n=13,976)	Control Percent (n=13,437)	Test minus Control	P-Value
92.4 (0.4)	90.5 (0.4)	1.9 (0.5)	<0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding.
P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test (test ≠ control) at the α=0.1 level.

Is the proportion of "Dial-up" internet service the same for test and control treatments?

Response proportions for the various types of internet subscriptions can be found in Table 18. Once again, please note that the universe for the subscriptions question is households that access the internet with a subscription. Starting with dial-up, we see that there is no significant difference in the share of households reporting this type of subscription in the test versus control treatments. This is as expected, given similar wording for this category in the two versions of the subscription type question.

Table 18. Response Proportions for Control and Test Treatments, Internet Subscription Type Question

Category	Test Percent (n=14,037)	Control Percent (n=13,476)	Test minus Control	Alternative Hypothesis	Adjusted P-Value
Dial-up	2.3 (0.2)	2.7 (0.2)	-0.4 (0.2)	T≠C	0.23
High speed vs. DSL/Cable/Fiber-optic	81.4 (0.5)	85.0 (0.5)	-3.6 (0.6)	T≠C	<0.01*
Cellular data plan vs. Mobile broadband	79.9 (0.4)	39.7 (0.6)	40.2 (0.8)	T>C	<0.01*
Satellite	6.5 (0.3)	6.0 (0.3)	0.5 (0.4)	T≠C	0.44
Other service	1.7 (0.2)	1.6 (0.1)	0.1 (0.2)	T≠C	0.61

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference between the two rates at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method. Question allows for multiple categories to be marked, so columns will not sum to 100 percent.

Is the proportion of "Yes" responses obtained by collapsing the control categories of "DSL," "Cable," and "Fiber-optic" the same as the proportion of "Yes" responses for the test treatment category of "Broadband (high speed)?"

The second row of Table 18 above displays the proportion of households reporting a broadband service such as DSL, cable, or fiber-optic. The share of households reporting this type of internet service is lower in the test treatment, at 81.4 percent, than in the control treatment, at 85.0 percent. While the difference is significant, the results are close to what we were expecting. This difference likely reflects the number of categories measuring this type of service. Respondents had three categories of this type in the control version of the question, but a single category in the test version. We are unable to determine whether the difference indicates overreporting for

the control version or underreporting for the test version. However, this does indicate an unintended consequence of streamlining the question.

Is the proportion of "Cellular data" higher in the test treatment than "Mobile broadband plan" is in the control?

Looking once more at Table 18 above, we see a striking result for the share of households reporting cellular or mobile internet service. Reports of this type of service are about twice as high in the test treatment, at 79.9 percent, compared with the control treatment, at 39.7 percent. This finding suggests that respondents understand the phrase "Cellular data plan" more clearly than "Mobile broadband plan." The movement of the category to the first position under the question stem in the test treatment may also have made the choice more visible to respondents.

Is the proportion of "Satellite" internet service the same for test and control treatments?

Results for satellite service in Table 18 indicate that there is no significant difference in the proportion of households reporting satellite internet service in the test versus control treatments. These results were expected as there was no change to the wording for satellite internet service category.

Is the proportion of "Some other service" in the test treatment less than or equal to the proportion in the control treatment?

The final row of Table 18 above contains results for the share of households reporting "some other service". There is no significant difference between the test treatment and the control treatment. This was expected due to the fact that there was no difference in question wording between control and test for this response category.

Among households that reported having a smartphone or tablet computer in the computers question, is the proportion reporting "Yes" to "Mobile broadband" higher in test than in control?

Finally, Table 19 displays the proportion of households reporting mobile broadband, focusing on households owning a device such as a smartphone or tablet. Similar to what was seen for all households, the share reporting mobile broadband is strikingly higher for households in the test treatment (85.4 percent) than those in the control treatment (43.3 percent). This result indicates that the new question wording improves measurement of mobile broadband not only for households overall, but also for those owning or using handheld devices.

Table 19. Proportion of Households with Smartphone or Tablet Reporting Mobile Broadband

Test Percent (n=12,758)	Control Percent (n=11,818)	Test minus Control	Alternative Hypothesis	P-Value
85.4 (0.4)	43.3 (0.7)	42.2 (0.8)	T>C	<0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. P-values with an asterisk (*) indicate a significant difference between the two rates at the α=0.1 level. Minor additive discrepancies are due to rounding.

5.4 Response Error

Are the measures of response reliability (GDR and IOI) for each computer type category better for the test treatment than for the control treatment?

Table 20 displays the Gross Difference Rates (GDRs) from the control and test treatments for each category of the types of computers question. The reliability of responses on ownership of a desktop or laptop is not significantly different between treatments. However, as expected, the test treatment shows greater reliability regarding both smartphone or tablet use and use of some other computer. Seven percent of answers on smartphone or tablet use are inconsistent between the original interview and CFU for the test treatment, whereas 10.8 percent of answers on handheld use are inconsistent in the control treatment. Inconsistency in reports of owning some other computer is lower in the test treatment, at 11.2 percent, than in the control treatment, at 19.0 percent. Greater reliability for the test treatment is likely due to the addition of a category clarifying how tablets should be classified, as well as a category allowing respondents to report ownership of a smartphone specifically. Under the old (control) wording, some respondents may have reported their smartphone or tablet using the "Handheld" category in one interview, and under the "other computer" category in the other interview.

Table 20. Gross Difference Rates (GDRs) for Control and Test Treatments, Types of Computers Question

Category	Test Sample Size	Test GDR Percent	Control Sample Size	Control GDR Percent	Test minus Control	Adjusted P-Value
Desktop or laptop	7,766	6.0 (0.4)	7,799	6.0 (0.4)	<0.1 (0.6)	0.94
Smartphone or tablet vs. Handheld	7,746	7.0 (0.4)	7,771	10.8 (0.5)	-3.8 (0.7)	<0.01*
Other computer	7,728	11.2 (0.5)	7,748	19.0 (0.6)	-7.9 (0.8)	<0.01*

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test (test ≠ control) at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Indexes of Inconsistency (IOIs) for each category of the types of computers item are found in Table 21. Similar to results from GDRs, the IOIs indicate greater reliability in the reporting of smartphones or tablets for the test treatment. The IOI estimate for the test treatment (23.8 percent) is significantly lower than that for the control treatment (32.9 percent). Inconsistency in reports of using a desktop or laptop is not significantly different in test versus control, exhibiting

a low value in each treatment. Nor is inconsistency in answers for the "other computer" category significantly different in the test treatment compared with the control treatment. Values are high across treatments. Once more, results suggest that the new "Tablet" and specific "Smartphone" categories increase reliability. High inconsistency in reports of other computers is likely due to the inherent vagueness of "other" response options.

Table 21. Indexes of Inconsistency (IOIs) for Control and Test Treatments, Types of Computers Question

Category	Test Sample Size	Test IOI Percent	Control Sample Size	Control IOI Percent	Test minus Control	Adjusted P-Value
Desktop or laptop	7,766	18.3 (1.2)	7,799	20.5 (1.5)	-2.1 (1.8)	0.46
Smartphone or tablet vs. Handheld	7,746	23.8 (1.3)	7,771	32.9 (1.5)	-9.0 (2.2)	<0.01*
Other computer	7,728	88.0 (2.4)	7,748	89.7 (2.3)	-1.6 (3.3)	0.63

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test (test ≠ control) at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Are the measures of response reliability (GDR and IOI) better for the test treatment than for the control treatment for the internet access question?

GDRs for the control and test versions of the internet access question are presented in Table 22. Inconsistency in reports of access with a subscription is lower in the test treatment (9.4 percent) than in the control treatment (11.3 percent). Reliability for access without a subscription also improves under the revised question wording. About 4.2 percent of responses on access without a subscription are inconsistent between the original interview and reinterview for the test treatment, compared with 9.1 percent of responses in the control treatment. Reliability for the "No internet access" category is not significantly different between test and control. Respondents likely interpret the term "paying," used in the test version of the question, in a more consistent way than the term "subscription," used in the control version. Also, adding the phrase "cell phone company" likely increases reliability for respondents who receive internet through a cell phone service instead of or in addition to a conventional internet service provider.

Table 22. Gross Difference Rates (GDRs) for Control and Test Treatments, Internet Access Question

Category	Test GDR Percent (n=7,669)	Control GDR Percent (n=7,641)	Test minus Control	Adjusted P-Value
Access with subscription	9.4 (0.5)	11.3 (0.5)	-2.0 (0.7)	<0.01*
Access without subscription	4.2 (0.4)	9.1 (0.5)	-4.9 (0.6)	<0.01*
No internet access	6.6 (0.4)	5.7 (0.4)	0.9 (0.6)	0.13

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding.
P-values with an asterisk (*) indicate a significant difference based a two-tailed t-test (test ≠ control) at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Table 23 contains the IOIs for each category of the internet access question, as well as the L-fold index of inconsistency (IOI_L) capturing reliability for the overall question. Starting with the IOI_L, we see that the reliability of estimates of internet access is not significantly different between test and control treatments. As a whole, the internet access question demonstrates moderate levels of inconsistency. Similarly, the IOIs for the individual access categories are not significantly different in the test treatment. Levels of inconsistency are moderate for the "access with a subscription" and "no internet access" categories, and high for the "access without a subscription" response option. The high inconsistency of the "access without a subscription" category likely relates to its status as a residual category. Because the legislation governing this topic in ACS specifies that internet subscriptions be measured, this response option is needed to make the internet access question exhaustive. However, respondents may interpret this category differently at various points in time. For example, respondents whose apartment building provides internet service could initially say that they have access without a subscription, since they do not directly subscribe. But at a later point, they could report access with a subscription, thinking that they do pay for internet through higher rent.

Table 23. Indexes of Inconsistency (IOIs) for Control and Test Treatments, Internet Access Question

Category	Test IOI Percent (n=7,669)	Control IOI Percent (n=7,641)	Test minus Control	Adjusted P-Value
Entire question (IOI_L)	34.9 (1.8)	39.5 (1.5)	-4.7 (2.5)	0.23
Access with subscription	33.4 (1.8)	36.1 (1.6)	-2.8 (2.6)	0.69
Access without subscription	71.8 (5.2)	78.9 (3.2)	-7.1 (6.0)	0.69
No internet access	27.6 (1.8)	24.6 (1.5)	3.0 (2.5)	0.69

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. Significance was tested based on a two-tailed t-test (test ≠ control) at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Are the measures of response reliability (GDR and IOI) for each internet subscription type better for the test treatment than for the control treatment?

Turning to the internet subscription type question, the GDRs for the control and test treatments are found in Table 24. As a final reminder, the universe for the subscription question is households that access the internet with a subscription. For "dial-up" and "other service" subscription types there are no significant differences between test and control. The GDR for the "Satellite" category is higher in the test than in the control treatment. In contrast, inconsistency between the original interview and reinterview is lower for the test item on cellular data plans (17.4 percent) than for the control item on mobile broadband (38.1 percent). Thus, respondents interpret the phrase "Cellular data plan" more consistently than the phrase "Mobile broadband."

Inconsistency is higher for the test version of high speed internet versus the control version of the combined categories of DSL, Cable, and Fiberoptic. The need to combine categories to make a straight comparison between treatments may have contributed to the lower gross difference rate for control as the probability of consistency is higher for three combined categories than for one category on its own.

Table 24. Gross Difference Rates (GDRs) for Control and Test Treatments, Internet Subscription Type Question

Category	Test Sample Size	Test GDR Percent	Control Sample Size	Control GDR Percent	Test minus Control	Adjusted P-Value
Dial-up	5,950	4.6 (0.4)	5,527	3.9 (0.4)	0.8 (0.6)	0.40
High speed vs. DSL/Cable/Fiber-optic	5,927	13.0 (0.8)	5,531	9.9 (0.5)	3.1 (0.9)	<0.01*
Cellular data plan vs. Mobile broadband	5,965	17.4 (0.8)	5,494	38.1 (1.0)	-20.8 (1.3)	<0.01*
Satellite	5,954	9.4 (0.7)	5,523	6.5 (0.4)	2.9 (0.8)	<0.01*
Other service	5,945	4.9 (0.5)	5,511	4.2 (0.5)	0.7 (0.6)	0.40

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test (test ≠ control) at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

Finally, Table 25 contains the IOIs for each internet subscription type. These results indicate that reliability for dial-up, broadband (high speed), satellite, or other service is not significantly different when comparing test versus control. Again, we find evidence of greater reliability for estimates of cellular data plans from the test treatment compared with estimates of mobile broadband from the control treatment. The IOI test estimate, at 52.3 percent, is significantly lower than the IOI control estimate, at 76.5 percent. In general, levels of inconsistency for the various subscription types are high, with index values over 50 percent. Once more, these findings suggest that respondents more reliably understand the phrase "Cellular data plan" than the phrase "Mobile broadband."

Table 25. Indexes of Inconsistency (IOIs) for Control and Test Treatments, Internet Subscription Type Question

Category	Test Sample Size	Test IOI Percent	Control Sample Size	Control IOI Percent	Test minus Control	Adjusted P-Value
Dial-up	5,950	85.5 (3.6)	5,527	85.4 (4.8)	0.1 (5.9)	1.00
High speed vs. DSL/Cable/Fiber-optic	5,927	53.2 (2.8)	5,531	53.2 (2.3)	-0.1 (3.3)	1.00
Cellular data plan vs. Mobile broadband	5,965	52.3 (2.1)	5,494	76.5 (2.1)	-24.2 (3.1)	<0.01*
Satellite	5,954	65.6 (3.4)	5,523	58.2 (2.9)	7.4 (3.9)	0.23
Other service	5,945	93.5 (4.3)	5,511	96.7 (2.0)	-3.2 (4.5)	1.00

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are shown in parentheses. Minor additive discrepancies are due to rounding. P-values marked with an asterisk (*) indicate a significant difference based on a two-tailed t-test (test ≠ control) at the α=0.1 level. P-values have been adjusted for multiple comparisons using the Holm-Bonferroni method.

6 CONCLUSIONS AND RECOMMENDATIONS

Questions on Computer and Internet Use were first introduced to the ACS in 2013. Considering the brisk rate at which technology develops and changes, question revisions were already needed. Specific concerns included the relatively low percentage of handheld-owning households reporting an internet subscription, as well as low reports of mobile broadband subscriptions (File & Ryan, 2014). The 2016 ACS Content tested several changes to the Computer and Internet Use questions. The primary change to the types of computers question involved the replacement of the "Handheld computer" category with a specific "Smartphone" category and a new category for "Tablet or other portable wireless computer." For the internet access question, the main changes involved replacing the term "subscription" with "paying," and asking about payment to a cell phone company in addition to an internet service provider. Substantial changes to the subscription type question involved replacing the phrase "Mobile broadband plan" with "Cellular data plan," and moving this category to the top position. In addition, the individual categories for "DSL," "Cable," and "Fiber-optic" were combined into a single "Broadband (high speed)" category.

Overall, results indicate that data quality improved when using the revised questions. All of the key research criteria for the internet access question were met, and four of five key criteria were met for both the types of computers and internet subscription type questions. In each case, the key criterion not met was the question of lowest priority.

Item missing data rates in the test treatment were either significantly lower than or not significantly different from those in the control treatment across the board. Revised wording showed improvements in nonresponse for the internet access question and for the "Cellular data plan" subscription category.

Results for the response proportions analysis, in general, were as expected. Particularly noteworthy is the substantial increase in the share of households reporting a cellular data plan in the test treatment versus a mobile broadband plan in the control treatment. Whether looking at all households or specifically at households with a smartphone or tablet (handheld in control), the test proportion is about double the control proportion.

Contrary to expectations, the share of households owning a desktop or laptop is lower in the test treatment compared with the control treatment. This is likely due to some owners of tablets (but not desktops or laptops) in the control treatment marking the "Desktop, laptop, netbook, or notebook computer" category, due to the lack of a specific category for tablets. The second unexpected result and unmet key research criteria involves the share of households reporting a DSL, cable, or fiber-optic subscription in the control treatment versus a broadband (high speed) subscription in the test treatment. This proportion is lower in the test treatment. Once more, there is a likely explanation for this difference, as respondents were offered three categories of this type in the control version of the question but only one category in the test version.

Finally, findings from the response error analysis indicate that, across most Computer and Internet Use items, reliability is better or not significantly different for the test treatment compared with the control treatment. Worth noting is reduced inconsistency for the new "Tablet" and specific "Smartphone" categories, compared with the old "Handheld" category. For internet

subscriptions, there was an improvement for the "cellular data plan" category in the test treatment as compared to the "mobile broadband" category in the control treatment. As measured by the GDR, however, response reliability was less favorable in the test treatment than the control treatment for the "satellite internet" category and for the "high speed" category when compared to the combined "DSL/Cable/Fiberoptic" category. Even though these contrasts were not significant when using the IOI as the measure of reliability, the significant differences found in the GDRs provides evidence that respondent confusion may still be a problem with the test version. Due to the large improvement in reliability for the "cellular data plan" category in the test treatment, along with the other improvements to reliability in the computer use and internet access questions, the evidence suggests that in general the test questions performed better in terms of consistency of responses.

Altogether, the 2016 ACS Content Test and analyses presented here validate the decision to implement the revised question wording on the 2016 production ACS. Whether considering item missing data rates, response proportions, or response error; in general data quality is not significantly different or is improved given changes to the questionnaire. Especially promising is the higher share of households that indicates owning a smartphone or tablet reporting an internet subscription, and much higher reports of mobile broadband subscriptions. The revised question wording will be reflected in the 2016 ACS data release, scheduled to begin in September 2017.

7 ACKNOWLEDGEMENTS

The 2016 ACS Content Test would not have been possible without the participation and assistance of many individuals from the Census Bureau and other agencies. Their contributions are sincerely appreciated and gratefully acknowledged.

- Census Bureau staff in the American Community Survey Office, Application Development and Services Division, Decennial Information Technology Division, Decennial Statistical Studies Division, Field Division, National Processing Center, Population Division, and Social, Economic, and Housing Statistics Division.
- Representatives from other agencies in the Federal statistical system serving on the Office of Management and Budget's Interagency Working Group for the ACS and the Topical Subcommittees formed by the Working Group for each topic tested on the 2016 ACS Content Test.
- Staff in the Office of Management and Budget's Statistical and Science Policy Office.

The authors would like to thank the following individuals for their contributions to the analysis and review of this report: Kurt Bauman, Nicole Scanniello, Jason Lee, Broderick Oliver, and Elizabeth Poehler.

8 REFERENCES

Biemer, P. (2011). *Latent Class Analysis of Survey Error*. Wiley, New York.

Dusch, G., & Meier, F. (2012). *2010 Census Content Reinterview Survey Evaluation Report.* U.S. Census Bureau, June 13, 2012. Retrieved January 13, 2017 from http://www.census.gov/2010census/pdf/2010_Census_Content_Reinterview_Survey_Evaluation_Report.pdf

File, T., & Ryan, C. (November 2014). *Computer and Internet Use in the United States: 2013.* U.S. Census Bureau. Retrieved January 13, 2017 from http://www.census.gov/content/dam/Census/library/publications/2014/acs/acs-28.pdf

Flanagan, P. (1996). *Survey Quality & Response Variance* (Unpublished Internal Document). U.S. Census Bureau. Demographic Statistical Methods Division. Quality Assurance and Evaluation Branch.

Flanagan, P. (2001). *Measurement Errors in Survey Response*. University of Maryland Baltimore County, Baltimore, Maryland.

Holm, S. (1979). "A Simple Sequentially Rejective Multiple Test Procedure," Scandinavian Journal of Statistics, Vol. 6, No. 2: 65-70. Retrieved January 11, 2017 from https://www.jstor.org/stable/4615733?seq=1#page_scan_tab_contents

National Telecommunications and Information Administration (October 2016). *Digital Nation Data Explorer.* Retreived June 12, 2017, from https://www.ntia.doc.gov/data/digital-nation-data-explorer#sel=internetUser&disp=map

Pew Research Center. (January 2014). *E-Reading Rises as Device Ownership Jumps.* Retrieved January 13, 2017 from http://www.pewinternet.org/files/2014/01/PIP_E-reading_011614.pdf

Pew Research Center. (October 2015). *Technology Device Ownership: 2015.* Retrieved December 29, 2016 from http://www.pewinternet.org/files/2015/10/PI_2015-10-29_device-ownership_FINAL.pdf

Pew Research Center. (December 2015). *Home Broadband 2015.* Retrieved December 29, 2016 from http://www.pewinternet.org/files/2015/12/Broadband-adoption-full.pdf

Pew Research Center. (September 2016). *Book Reading 2016.* Retrieved December 29, 2016 from http://assets.pewresearch.org/wp-content/uploads/sites/14/2016/08/PI_2016.09.01_Book-Reading_FINAL.pdf

Pew Research Center. (November 2016). *Device ownership 2016.* Retrieve June 12, 2017 from http://www.pewresearch.org/data-trend/media-and-technology/device-ownership/

Rao, J. N. K.; Scott, A. J. (1987). "On Simple Adjustments to Chi-Square Tests with Sample Survey Data," The Annal of Statistics, Vol. 15, No. 1, 385-397. Retrieved on January 31, 2017 from http://projecteuclid.org/euclid.aos/1176350273

Reichert, J. (2015, July 7). *2016 Mandated Content Changes.* U.S. Census Bureau.

Stapleton, M., & Steiger, D. (2015). *Cognitive Testing of the 2016 American Community Survey Content Test Items: Summary Report for Round 1 and Round 2 Interviews.* Westat, Rockville, Maryland, January 2015.

Steiger, D., Anderson, J., Folz, J., Leonard, M., & Stapleton, M. (2015). *Cognitive Testing of the 2016 American Community Survey Content Test Items: Briefing Report for Round 3 Interviews.* Westat, Rockville, Maryland, June 2015.

U.S. Census Bureau. (2014). *American Community Survey Design and Methodology (January 2014).* Retrieved January 13, 2017 from http://www2.census.gov/programssurveys/acs/methodology/design_and_methodology/acs_design_methodology_report_2014.pdf

U.S. Census Bureau (2016). *2015 Planning Database Tract Data* [Data file]. Retrieved January 11, 2017 from http://www.census.gov/research/data/planning_database/2015/

Appendix A. Control and Test Questions in CATI, CAPI, and CFU

Figure A1. CATI/CFU and CAPI Versions of the Control and Test Questions

Control Version	Test Version
[LAPTOP] **For the next few questions about computers, EXCLUDE GPS devices, digital music players, and devices with only limited computing capabilities, for example: household appliances.** **At this <house/apartment/mobile home/unit>, do you or any member of this household own or use a desktop, laptop, netbook, or notebook computer?** Yes No **[HANDHELD]** **At this <house/apartment/mobile home/unit>, Do you or any member of this household own or use a handheld computer, smart mobile phone, or other handheld wireless computer?** Yes No **[COMPOTH]** **At this <house/apartment/mobile home/unit>, Do you or any member of this household own or use some other type of computer?** Yes No (Skip to internet access question) **[COMPOTHW]** **What is this other type of computer? _____** --- **[WEB]** **At this <house/apartment/mobile home/unit>, do you or any member of this household access the Internet?** Yes No (Skip to vehicle question) **[SUBSCRIBE]** **At this <house/apartment/mobile home/unit>, Do you or any member of this household access the Internet with or without a subscription to an Internet service?** With a subscription to an Internet service Without a subscription to an Internet service (Skip to vehicle question) --- **[DIALUP]** **At this <house/apartment/mobile home/unit>, do you or any member of this household subscribe to the Internet using a dial-up service?** Yes No	**[LAPTOP]** **At this <house/apartment/mobile home/unit>, do you or any member of this household own or use a desktop or laptop computer?** Yes No **[SMARTPHONE]** **At this <house/apartment/mobile home/unit>, Do you or any member of this household own or use a smartphone?** Yes No **[TABLET]** **At this <house/apartment/mobile home/unit>, Do you or any member of this household own or use a tablet or other portable wireless computer?** Yes No **[COMPOTH]** **At this <house/apartment/mobile home/unit>, Do you or any member of this household own or use some other type of computer?** Yes No (Skip to Internet access question) **[COMPOTHW]** **What is this other type of computer? _____** --- ****[ACCESS] – Internet mode** ****[WEB] – CATI/CAPI/CFU** **At this <house/apartment/mobile home/unit>, do you or any member of this household have access to the Internet?** Yes No (Skip to vehicle question) **[SUBSCRIBE]** **At this <house/apartment/mobile home/unit>, Do you or any member of this household pay a cell phone company or Internet service provider to access the Internet?** Yes No (Skip to vehicle question) ---

Figure A1. (continued). CATI/CFU and CAPI Versions of the Control and Test Questions

Control Version	Test Version
[DSL] **At this <house/apartment/mobile home/unit>, Do you or any member of this household subscribe to the Internet using a DSL service?** Yes No **[MODEM]** **At this <house/apartment/mobile home/unit>, Do you or any member of this household subscribe to the Internet using a cable-modem service?** Yes No **[FIBEROP]** **At this <house/apartment/mobile home/unit>, Do you or any member of this household subscribe to the Internet using a fiber-optic service?** Yes No **[BROADBND]** **At this <house/apartment/mobile home/unit>, Do you or any member of this household subscribe to the Internet using a mobile broadband plan for a computer or a cell phone?** Yes No **[SATELLITE]** **At this <house/apartment/mobile home/unit>, Do you or any member of this household subscribe to the Internet using a satellite Internet service?** Yes No **[OTHSVCE]** **At this <house/apartment/mobile home/unit>, Do you or any member of this household subscribe to the Internet using some other service?** Yes No (Skip to vehicle question) **[OTHSVCEW]** **What is this other type of Internet service? _____**	**[BROADBND]** **Do you or any member of this household access the Internet using a cellular data plan for a smartphone or other mobile device?** Yes No **[HISPEED]** **Do you or any member of this household access the Internet using broadband or high speed Internet service such as cable, fiber optic, or DSL service installed in this <house/apartment/mobile home/unit>?** Yes No **[SATELLITE]** **Do you or any member of this household access the Internet using a satellite Internet service installed in this <house/apartment/mobile home/unit>?** Yes No **[DIALUP]** **Do you or any member of this household access the Internet using a dial-up Internet service installed in this <house/apartment/mobile home/unit>?** Yes No **[OTHSVCE]** **Do you or any member of this household access the Internet using some other service?** Yes No (Skip to vehicle question) **[OTHSVCEW]** **What is this other type of Internet service?** _____

Appendix B. Unit Response Rates Supplemental Table

Table B1. Unit Response Rates by Designated High (HRA) and Low (LRA) Response Areas

Mode	Test Interviews	Test Percent	Control Interviews	Control Percent	Test minus Control	P-Value
Total Response	19,400		19,455			
HRA	7,556	94.3 (0.4)	7,608	94.5 (0.3)	-0.2 (0.6)	0.72
LRA	11,844	91.5 (0.3)	11,847	91.0 (0.3)	0.5 (0.5)	0.29
Difference		2.7 (0.5)		3.5 (0.5)	-0.7 (0.7)	0.33
Self-Response	13,131		13,284			
HRA	6,201	59.7 (0.7)	6,272	60.6 (0.7)	-0.9 (0.9)	0.31
LRA	6,930	33.2 (0.4)	7,012	33.6 (0.4)	-0.4 (0.6)	0.55
Difference		26.5 (0.8)		27.0 (0.8)	-0.5 (1.2)	0.66
Internet	8,168		8,112			
HRA	4,119	39.6 (0.6)	4,048	39.1 (0.6)	0.5 (0.8)	0.51
LRA	4,049	19.4 (0.3)	4,064	19.5 (0.3)	0.1 (0.4)	0.87
Difference		20.2 (0.6)		19.6 (0.7)	0.6 (0.9)	0.52
Mail	4,963		5,172			
HRA	2,082	20.0 (0.4)	2,224	21.5 (0.4)	-1.5 (0.6)	0.02*
LRA	2,881	13.8 (0.3)	2,948	14.1 (0.3)	-0.3 (0.4)	0.43
Difference		6.2 (0.5)		7.4 (0.4)	-1.1 (0.7)	0.11
CATI	872		880			
HRA	296	9.0 (0.5)	301	9.6 (0.6)	-0.6 (0.8)	0.44
LRA	576	7.9 (0.4)	579	8.0 (0.3)	-0.1 (0.5)	0.85
Difference		1.1 (0.6)		1.6 (0.7)	-0.5 (0.9)	0.58
CAPI	5,397		5,291			
HRA	1,059	82.2 (1.0)	1,035	82.7 (0.9)	-0.5 (1.3)	0.69
LRA	4,338	85.8 (0.5)	4,256	85.0 (0.4)	0.8 (0.7)	0.23
Difference		-3.7 (1.1)		-2.3 (1.0)	-1.3 (1.5)	0.36

Source: U.S. Census Bureau, 2016 American Community Survey Content Test
Note: Standard errors are in parentheses. Minor additive discrepancies are due to rounding. P-values with an asterisk (*) indicate a significant difference based on a two-tailed t-test at the α=0.1 level. The weighted response rates account for initial sample design as well as CAPI subsampling.

Appendix C. Benchmarks

C.1. Research Questions

1. How do the proportions for each category of computers in each treatment compare with proportions found in the Current Population Survey (CPS) and from surveys done by the Pew Research Center?
2. How do the proportions in each treatment compare with proportions found in the CPS for the Internet access question?
3. How do the proportions of mobile broadband subscribers compare to Pew Research findings as well as the most recent CPS results?

C.2. Methodology

We compared the 2016 ACS Content Test data from both control and test treatments with the most current version of other surveys available as benchmarks for the comparisons. These comparisons allow us to tell whether our results differ from other reliable sources. As a cautionary note, although the other surveys provide benchmarks, they are not statistically comparable with the Content Test results, given differences in universe, timing, question wording, and survey design between the sources. Useful comparisons can still be made, however, as the overall distributions should be similar between surveys or differ in expected ways.

Types of Computers

For the topic of types of computers, we compared data from both control and test treatments to information from the July 2015 Current Population Survey (CPS) Computer and Internet Use Supplement and recent Pew Research Center surveys.

The CPS, a national household survey, has collected data on computer use since 1984 and internet use since 1997 in an occasional supplement. The July 2015 CPS Supplement included questions about access to desktops, laptops, smartphones, and tablets, as well as wearable technology and smart TVs.[20] For our comparison, we looked at CPS estimates on use of 1) a desktop or laptop/notebook, 2) a smartphone, and 3) a tablet/e-book reader or a wearable internet-connected device (such as a smart watch or glasses, with the item offering specific examples). The smartphone estimate is a recode rather than a direct question, as the direct question on the survey asks about both cellular phones and smartphones. Following guidance by the National Telecommunications and Information Administration (NTIA), the supplement sponsor, we created the smartphone recode using the item on cellular phones and smartphones, combined with information on internet use from any location and subscription to a mobile data plan. Similar to the Content Test, the universe for the CPS is households.

The Pew Research Center began asking about cellphone ownership in 2000, desktops or laptops in 2004, tablets in 2010, and smartphones in 2011. The most recent data on smartphones is from 2015, and 2016 data on laptops/desktops and tablets are available. Pew respondents receive

[20] Complete technical documentation, including question wording, for the 2015 CPS Computer and Internet Use Supplement is available at http://www2.census.gov/programs-surveys/cps/techdocs/cpsjul15.pdf.

direct questions about whether they have 1) a desktop or laptop computer, and 2) a tablet (with the item offering specific examples). Note that question wording regarding tablet ownership in the Pew survey differs somewhat from the Content Test wording. The test treatment asks about owning or using a "Tablet or other portable wireless computer," whereas the comparable Pew item only asks about tablet computers. Smartphone owners are identified through two questions, with the first asking about having a cell phone, and the second asking if the person's cell phone is a smartphone (with the item noting examples).[21] Pew data are typically collected for adults and are therefore not statistically comparable with ACS data on Computer and Internet Use, which represent results for each housing unit (Pew Research Center, 2015).

Proportions of desktop/laptop ownership from both the test and control treatments were used in comparisons. However, only test proportions on smartphone and tablet ownership were used, as the control treatment lacks categories specific to these devices.

Internet Access

For the topic of internet access, we again compared data from both control and test treatments to information from the 2015 CPS Supplement and a recent Pew Research Center survey.

The 2015 CPS Supplement asked a series of five questions about how household members connect to the internet at home. Those who stated they used a plan bought from 1) a company, or 2) a public agency, nonprofit, or cooperative were considered to have access with a subscription. As for the Content Test, the universe for the CPS estimate is households.

The Pew estimate for access with a subscription was captured in 2015, and includes those with either a smartphone or a home broadband subscription. Pew separately reported on dial-up subscribers. As noted above, Pew estimates have a universe of adults, in contrast with the Content Test universe of households.

Internet Subscription

For the topic of internet subscription type, we compared data from both control and test treatments to information from the 2015 CPS Supplement. Although most internet subscription types are not captured in Pew Research Center surveys, we were able to use a Pew estimate on dial-up service for comparison.

The 2015 CPS questions contain categories similar to the ACS Content Test: 1) mobile internet service or a data plan; 2) high-speed internet such as cable, DSL, or fiber-optic; 3) satellite; 4) dial-up; and 5) some other service. Also similar to the Content Test, the universe for CPS estimates is households with an internet subscription. As with other Pew estimates, the universe for the Pew dial-up estimate is adults aged 18 and over.

[21] Pew collects data through phone interviews, sampling both those with landline phones and those with cellular phones. The cellular phone sample is automatically marked as having a cell phone (Pew Research Center, October 2015).

C.3. Results

How do the proportions for each category of computers in each treatment compare with proportions found in the Current Population Survey (CPS) and from surveys done by the Pew Research Center?

Table C1 contains proportions of households owning various types of computers from the test and control treatments. It also contains proportions from the 2015 CPS Computer and Internet Use Supplement and recent surveys by the Pew Research Center. Although CPS and Pew estimates are not statistically comparable to those from the Content Test, useful comparisons can still be made. An important difference to note across surveys is that the Content Test and CPS estimates are for households, whereas Pew estimates are for adults aged 18 and over. Because not all household members in a household with a given type of computer would be expected to report that they own such a computer, the percentage of households with that type of computer should be larger than the percentage of people with that type. In addition, because computer ownership and use have been growing, data collected at a later time would be expected to show higher levels of ownership than those collected at an earlier time.

Table C1. Benchmark Estimates, Types of Computer Question

Item	Test Percent	Control Percent	CPS Percent	Pew Percent
Desktop or laptop	79 (± 0.7)	81 (± 0.7)	71 (± 0.3)	74 (± 2.4)
Smartphone	78 (± 0.7)	N/A	62 (± 0.4)	68 (± 2.1)
Tablet	60 (± 0.8)	N/A	39 (± 0.4)	48 (± 2.4)

Sources: U.S. Census Bureau, 2016 American Community Survey Content Test and 2015 Current Population Survey Computer and Internet Use Supplement; Pew Research Center (Surveys conducted June 10-July 12, 2015 and March 7-April 4, 2016.)
Note: N/A indicates not applicable. Ninety percent margins of error are shown in parentheses. Estimates across surveys are not statistically comparable. Content Test and CPS estimates are for households, whereas Pew estimates are for adults aged 18 and over.

In general, we see that Content Test estimates of computer ownership—including desktop/laptop, smartphone, and tablet ownership—conform to expectations, with the Content Test showing results that, at surface value, are somewhat higher than Pew and CPS benchmarks. The large difference between CPS and Content Test estimates of tablet use is not easily explained. Nonetheless, these results suggest that the Content Test questions do a reasonably good job of measuring computer ownership.

How do the proportions in each treatment compare with proportions found in the Current Population Survey for the Internet access question?

Estimates of internet access from the test treatment, control treatment, CPS, and Pew are displayed in Table C2. In summary, Content Test proportions of access with a subscription conform to expectations relative to Pew estimates, adding to our confidence in Content Test estimates of internet access. The apparent difference between Content Test and CPS estimates of internet subscriptions is more problematic. This difference may partially result from CPS issues,

in that those data show a (nonsignificant) decline in household internet use between 2012 and 2015 not evident in other data.[22] This difference may be addressed in future research.

Table C2. Benchmark Estimates, Internet Access Question

Item	Test Percent	Control Percent	CPS Percent	Pew Percent
Access with subscription	84 (± 0.7)	82 (± 0.7)	71 (± 0.3)	80 (± 1.1)
No subscription	16 (± 0.7)	18 (± 0.7)	29 (± 0.3)	N/A

Sources: U.S. Census Bureau, 2016 American Community Survey Content Test and 2015 Current Population Survey Computer and Internet Use Supplement; Pew Research Center (Surveys conducted in April, July, and November 2015.)
Note: N/A indicates not applicable. Ninety percent margins of error are shown in parentheses. Estimates across surveys are not statistically comparable. Content Test and CPS estimates are for households, whereas Pew estimates are for adults aged 18 and over. Pew estimate represents those with a smartphone or home broadband connection. Another two percent have a dial-up connection.

How do the proportions of mobile broadband subscribers compare to Pew Research findings as well as the most recent CPS results?

Table C3 shows estimates of internet subscription type from the test treatment, control treatment, CPS supplement, and Pew. Again, whereas the universe for the types of computers and internet access questions is all eligible households, the Content Test universe for the internet subscriptions question is households with internet access with a subscription. Proportions of dial-up are not that different among the Content Test treatments, CPS, and Pew. Nor are differences great between the test treatment, control treatment, and CPS regarding high speed broadband service such as DSL, cable, or fiber-optic; satellite subscriptions; or subscription to some other service. Benchmarking the test and control estimates for households with a mobile broadband subscription against the CPS provides additional evidence that the new question wording improves measurement of mobile broadband. However, the difference between the test and CPS estimates for mobile broadband should be explored further in future research.

Table C3. Benchmark Estimates, Internet Subscription Type Question

Item	Test Percent	Control Percent	CPS Percent	Pew Percent
Dial-up	2 (± 0.3)	3 (± 0.3)	1 (± 0.1)	2 (± 2.3)
High speed or DSL/Cable/Fiber-optic	81 (± 0.8)	85 (± 0.8)	76 (± 0.4)	N/A
Cellular data plan or Mobile broadband	80 (± 0.7)	40 (± 1.0)	61 (± 0.4)	N/A
Satellite	6 (± 0.4)	6 (± 0.5)	3 (± 0.2)	N/A
Other service	2 (± 0.3)	2 (± 0.2)	1 (± 0.1)	N/A

Sources: U.S. Census Bureau, 2016 American Community Survey Content Test and 2015 Current Population Survey Computer and Internet Use Supplement; Pew Research Center (Survey conducted June 10-July 12, 2015.)
Note: N/A indicates not applicable. Ninety percent margins of error are shown in parentheses. Estimates across surveys are not statistically comparable. Content Test and CPS estimates are for households with an internet subscription, whereas Pew estimates are for adults aged 18 and over.

[22] See the relevant chart from NTIA's Digital Nation Data Explorer, available at https://www.ntia.doc.gov/data/digital-nation-data-explorer#sel=internetAtHome&demo=&pc=prop&disp=chart.