Cleared for Public Release

A Usability Evaluation of the PBA Search/NAICS Classification for the 2022 Economic Census Instrument

Submitted to:

Michelle Karlsson

Theresa Riddle

James Jamski

[1]Submitted by:

Temika Holland

Aryn Hernandez

Data Collection and Methodology Research Branch

January 15, 2021

Updated: March 23, 2021

Note: The information in this report is proprietary to the U. S. Census Bureau Data Collection and Methodology Research Branch. It may be used only to communicate findings to the organizational unit requesting the study and for professional research purposes.

The Census Bureau has reviewed this report for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied. (Approval ID: CBDRB-FY21-ESMD002-019)

[1] Report format based on ISO/IEC 25062:2006 Common Industry Format for Usability Test Reports

Table of Contents

1. INTRODUCTION 4
2. METHOD 4
2.1. PARTICIPANTS 4
2.2. STUDY DESIGN 6
2.3. TASKS 6
2.4. PROCEDURES 7
2.5. USABILITY METRICS AND DATA SCORING 7
3. RESULTS 8
3.1. Round 1: Findings and Recommendations 8
3.2. Round 2 Findings and Recommendations 29
4. Additional findings across rounds of testing: 33
4.1. SATISFACTION 35
APPENDIX A: Testing Protocol 38
APPENDIX B: User tasks for Round 1 testing 48

EXECUTIVE SUMMARY

The Data Collection and Methodology Research Branch (ESMD) performed testing on design features for the upcoming 2022 Economic Census online instrument between August and November 2020. The purposes of the usability evaluation was to 1) evaluate proposed changes in design as a result of the 2017 Economic Census evaluation (2) evaluate the Principal Business Activity (PBA)/ NAICS Classification functionality, (3) identify areas of the instrument that are problematic (4) identify instructions/features that are difficult to understand, and (5) provide recommendations for improvements to the design of the instrument.

Participants consisted of persons from various establishments located in the United States. The participants recruited were persons who responded to the 2017 Economic Census survey, the 2019 Annual Survey of Manufacturers and/or the Report of Organization Survey (ASM/COS) for the respective establishment. Participants completed simulated, but representative tasks that would be typically conducted during the response process for the Economic Census. These include adding/deleting locations, locating/identifying submission details, etc. A complete list of tasks used during testing can be found in Appendix B. Participants also interacted with a proposed NAICS search tool (incorporating machine learning) to help identify their principal business or activity.

During the 30-minute, one-on-one remote usability session, each participant was greeted by the Test Administrator (TA), informed about the purpose of the evaluation, and asked to review and sign a consent form. The TA instructed the participant to complete a set of tasks using the Economic Census prototypes and the NAICS tool via an external site, one at a time on a PC. The TA only intervened during the session when the participant experienced technical difficulties. During the session, the TA recorded the participants interaction with the materials and administered probing questions where necessary. At the end of the session the participants completed a satisfaction questionnaire and answered debriefing questions about their overall experience.

1. INTRODUCTION

At the request of the managers in the Economic Management Division's (EMD) Collection Instruments Branch, the Data Collection and Methodology Research Branch conducted a usability evaluation of various design features including the NAICS Classification tool for the upcoming 2022 Economic Census online instrument.

The Economic Census is conducted every 5 years by the U.S. Census Bureau and is a leading source of information about U.S. business and economy.

The following goals/objectives were identified for testing:

- Evaluate proposed changes in design resulting from the 2017 Economic Census evaluation
- Evaluate the Principal Business Activity (PBA)/ NAICS Classification functionality
- Identify areas of the instrument that are problematic
- Identify instructions/features that are difficult to understand
- Provide recommendations for improvements to the design of the instrument.

2. METHOD

2.1. PARTICIPANTS

Fifty-three participants took part across two rounds of usability testing for the 2022 Economic Census instrument (Round 1=29 participants; Round 2=24 participants). Participants had characteristics of potential respondents to the survey and varied in age and educational background. All participants reported experience with using the Internet.

Participants were recruited from single-unit and multi-unit establishments who reported to the 2019 Annual Survey of Manufactures/Report of Organization and/or the 2017 Economic Census survey. Participants held positions such as Controller, CFO, Accountant, etc. Their tenure in these positions varied from a few months to several years. The industries for which these participants represented varied and were reflective of the types of establishments reporting to the Economic Census. These included Manufacturing, Wholesale Trade, Construction, Retail Trade, Mining, and Other Services. See Table 1 and Table 2 for an overview of the number of participants by industry and size by round.

Table 1. Round 1: Establishments by Sector

Industry	Total
Service	7
Construction	-
Wholesale	8
Mining	-
Manufacturing	5
Retail	9
TOTAL	29

Table 2. Round 2: Establishments by Industry and Size

Industry	Total
Service	4
Construction	4
Wholesale	3
Mining	6
Manufacturing	4
Retail	3
TOTAL	24

2.2. STUDY DESIGN

The objectives of the usability evaluation were to (1) evaluate proposed changes in design as a result of the 2017 Economic Census evaluation (2) evaluate the Principal Business Activity (PBA)/ NAICS Classification functionality, (3) identify areas of the instrument that are problematic (4) identify instructions/features that are difficult to understand, and (5) provide recommendations for improvements to the design of the instrument. Findings derived from the evaluation can serve as a baseline for future iterations of the instrument as a way to benchmark current usability and identify areas where further improvements could be made. It may also be used to help inform decisions regarding the design and implementation of features/functions proposed for the upcoming 2022 Economic Census instrument.

While the objectives outlined above include an assessment of proposed designed changes from the 2017 Economic Census evaluation, another primary objective of testing was to evaluate the added functionality for NAICS Classification (2). The proposed feature would provide respondents the ability to "dynamically" determine the "correct" Economic Census questionnaire path, in real time, based on self-identifying their PBA. Respondents would be given the ability to enter a text description of their business and search for their PBA, if not prelisted. The hope is that the search functionality would aid in classification, ultimately, reducing the number of write in responses received. Additionally, screens surrounding this functionality (e.g., prescreens) were also included in the evaluation to add context and simulate how participants may encounter the NAICS search feature in a real survey environment.

During the usability tests, participants interacted with prototypes of design features proposed for the 2022 Economic Census instrument, and a test version of the NAICS Classification tool all accessed from the Qualtrics survey design platform. Testing occurred remotely using Skype for Business (when possible). Skype allowed for both audio and screen sharing capabilities, allowing the Researcher to see how the participant interacted with proposed design features. All participants received the same instructions during test administration (see Appendix A for a copy of the testing protocol).

2.3. TASKS

Round 1 participants' interacted with design features proposed for the NAICS Classification survey item, identifying the PBA for their respective establishment. Additionally, they were given tasks designed to reflect realistic and representative tasks that they would have to complete while reporting to the Economic Census. These tasks stemmed from prior research recommendation from an evaluation of the 2017 Economic Census instrument and consisted of the following:

- Adding locations
- Deleting locations
- Confirming Survey Submissions
- Using the Total by EIN feature

Round 2 testing did not include tasks associated with prior research of the 2017 Economic Census instrument (i.e., adding/deleting locations, confirming survey submissions, using the

Total by EIN feature). Round 2 of testing focused on assessing changes made to the design of features of the NAICS Classification tool, as a result of Round 1 testing.

2.4. PROCEDURES

Each usability session was conducted at the participants' business establishment. Sessions lasted approximately 30 minutes. All participants electronically signed a consent form referencing the OMB control number for this study, the confidentiality of the session, the volunteer nature of the study, and that the session will be recorded. The participant also completed a questionnaire about his/her Computer use and Internet experience.

The test administrator (TA) explained the "think-aloud" procedure in which participants are asked to verbalize their thoughts and behaviors as they interact with the instrument to provide a deeper understanding of their cognitive processes.

During Round 1 testing, participants were then asked to complete tasks (see Appendix B) using the instrument intended to evaluate design features from the 2017 Economic Census evaluation. Probes were administered as needed by the TA. Participants were shown each task in the Qualtrics survey design platform and asked to click on the prototype where they would expect to go to complete the task, or the information they would reference. For each task, participants were assigned a unique version of the prototype. After completing all tasks, participants interacted with the NAICS Classification tool. Lastly, participants were asked to answer a post-test satisfaction questionnaire about their overall experience using the features they encountered during testing. After this, the TA asked a set of debriefing questions and concluded the session.

Round 2 testing did not include an assessment of design features from the 2017 Economic Census evaluation, instead participants only interacted with the revised NAICS functionality as a result of findings from Round 1 testing- including prescreens surrounding the tool.

2.5. USABILITY METRICS AND DATA SCORING

The following table (Table 3) outlines performance metrics used to assess the usability of design features.

Table 3 Performance data scoring metric

Measures	Scoring
Effectiveness (accuracy) *Heatmap*	The heatmap depicts behavioral data outlining clicks for areas of the prototype instrument participants referenced to complete a given task. The spectrum ranges from green to red and is aggregated across participants. A higher propensity of clicks is indicated by red. Green indicates a lower propensity of clicks.

Subjective Satisfaction *Ratings*	Satisfaction ratings were based on subjective satisfaction ratings from the System Usability Scale questionnaire administered at the end of the usability session.

3. RESULTS

The findings from the usability tests were obtained according to the metrics specified in Section 2.5 and TA observations. Usability results are presented by round of testing.

3.1. Round 1: Findings and Recommendations

3.1.1. Adding locations

Task: Your company has gone through some changes and you need to make some updates to the location listing. Your company has new locations. You need to update your listing to reflect these new locations that aren't listed.
Click on where you would go to do this.

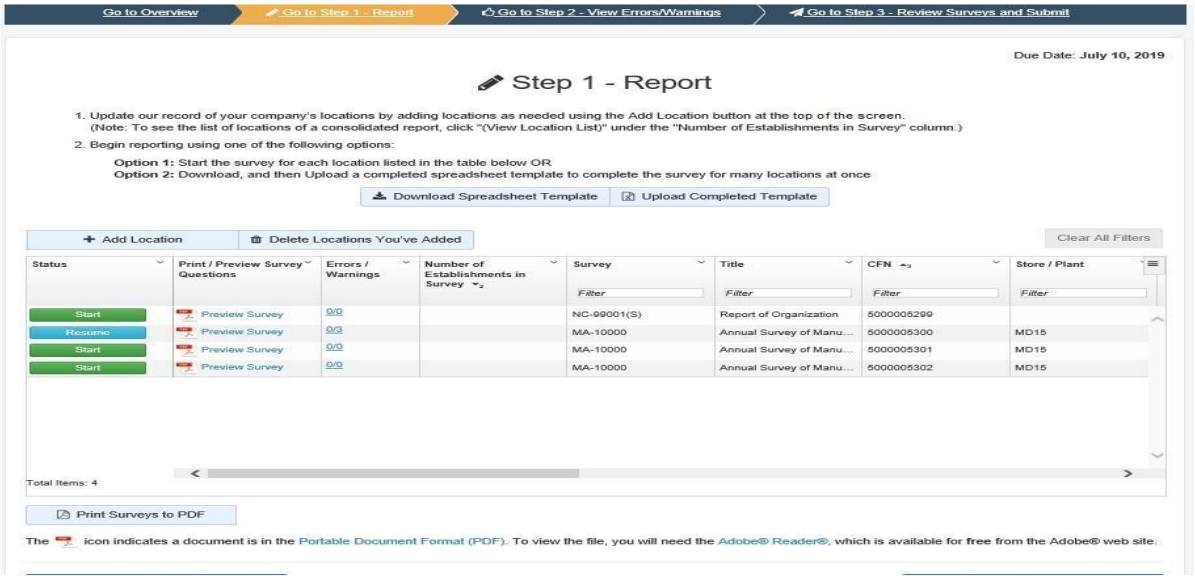


Figure 1. Proposed display with Add Location functionality at the top of the screen

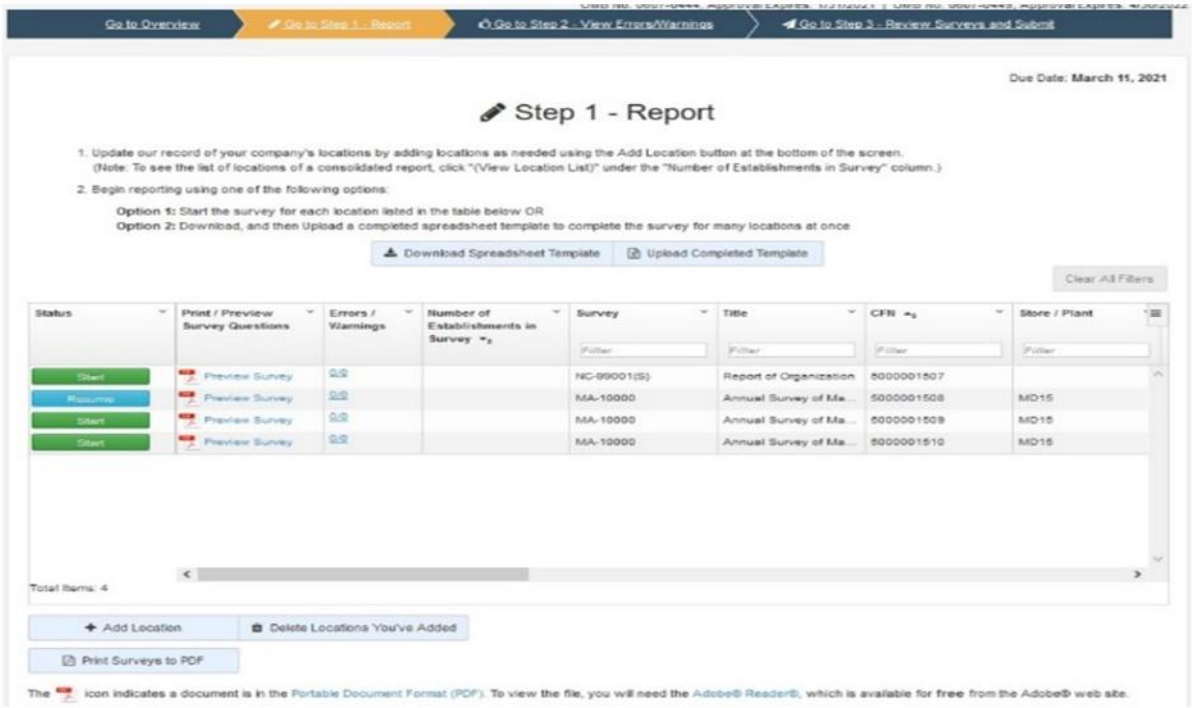


Figure 2. Proposed display with Add Location functionality at the bottom of the screen

During testing, participants were able to identify and click on the *Add locations* button needed to successfully complete the task whether it was at the bottom or top of the screen (as shown in the heatmap displayed in and)

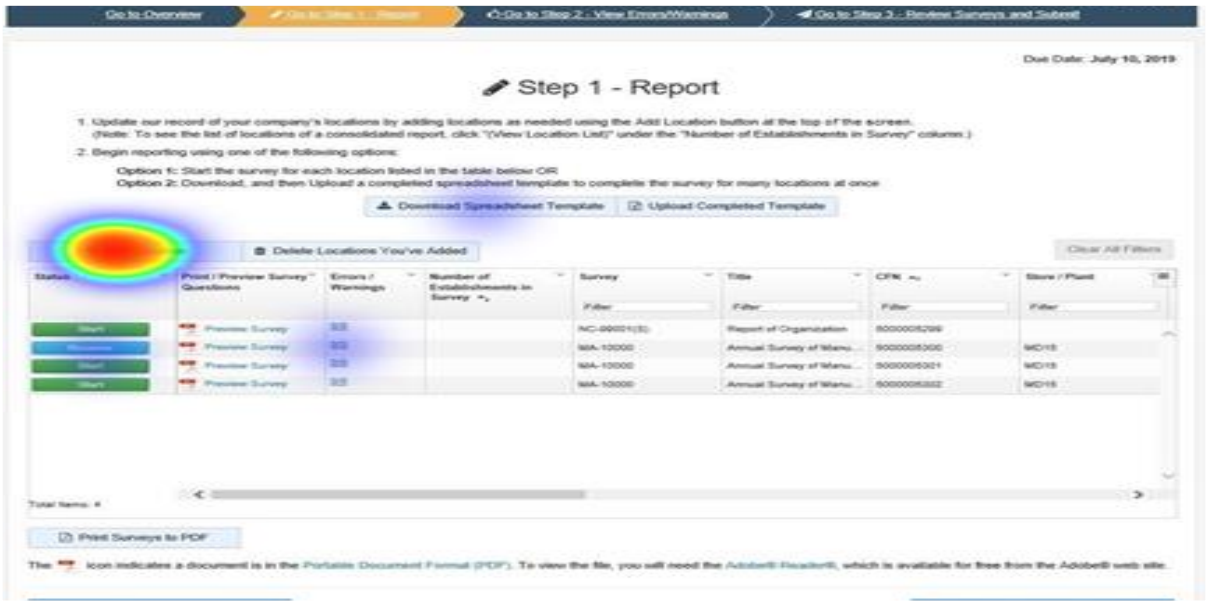


Figure 3. Behavioral data for *Adding locations* at the top of the screen design

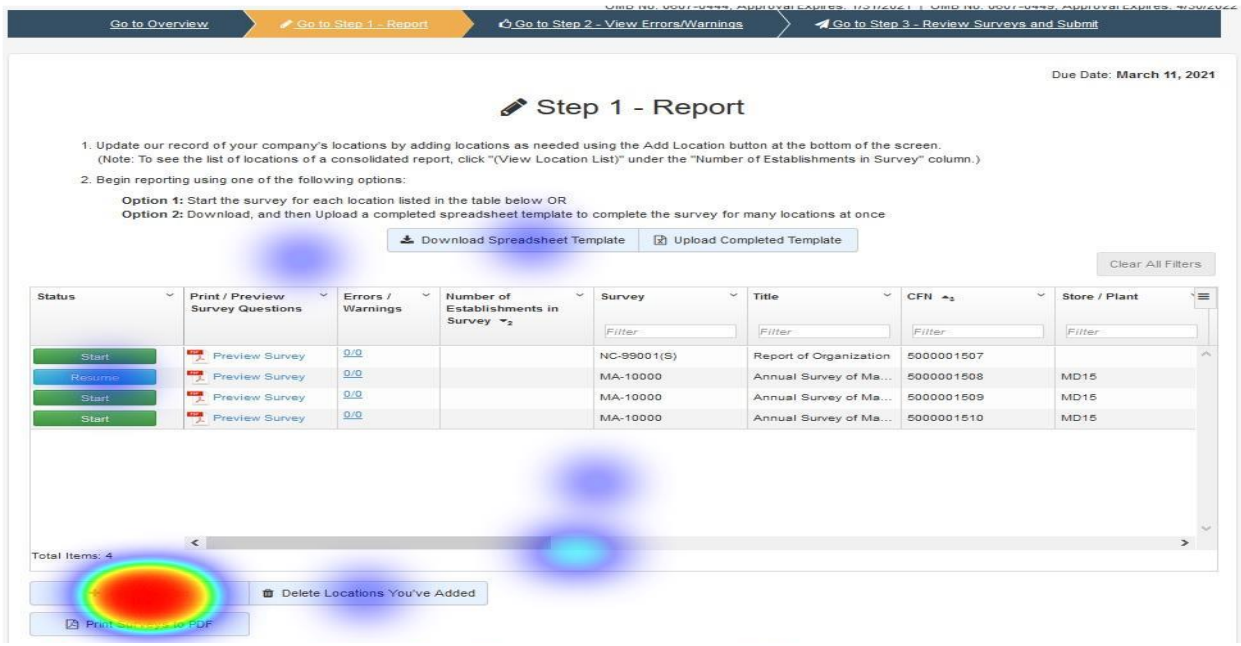


Figure 4. Behavioral data for *Adding locations* at the bottom of the screen design

Recommendation(s):

Although participants were successful in identifying the *Add Location* button to complete the task, it is recommended that the button be placed at the bottom of the screen (as shown in Version 2). Placing the button at the bottom of the screen may facilitate appropriate interaction with screen (i.e., review of prelisted location prior to making modifications to the listing)

Additionally, the inclusion of more buttons at the top of the screen may introduce clutter to the screen with the existing buttons already placed at the top (i.e., download and upload spreadsheet functionality)

3.1.2. Deleting locations

Task: The locations you added in the prior question no longer exist. You want to remove the locations from the listing. Is it clear which location(s) you are able to remove? Click on where you would go to remove them.

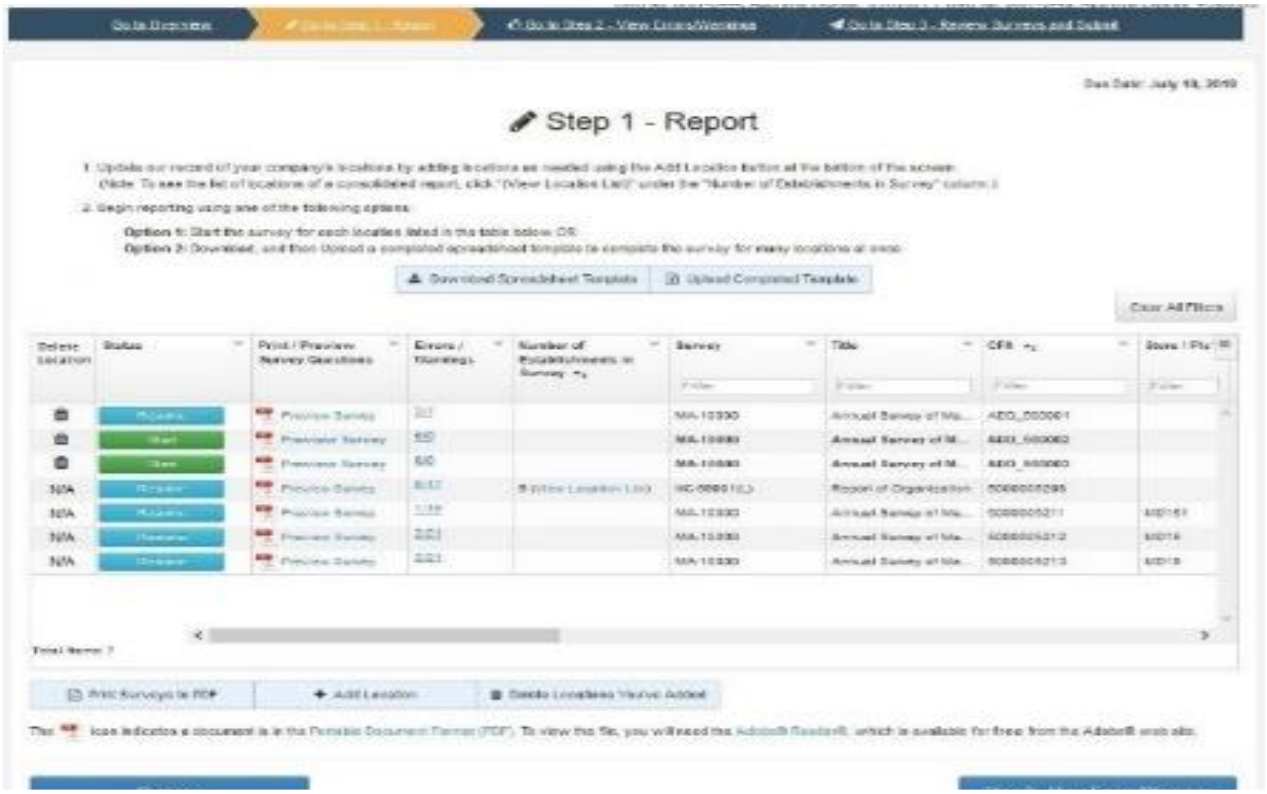


Figure 5. Proposed display for *Deleting locations* with trash can icon and N/A design

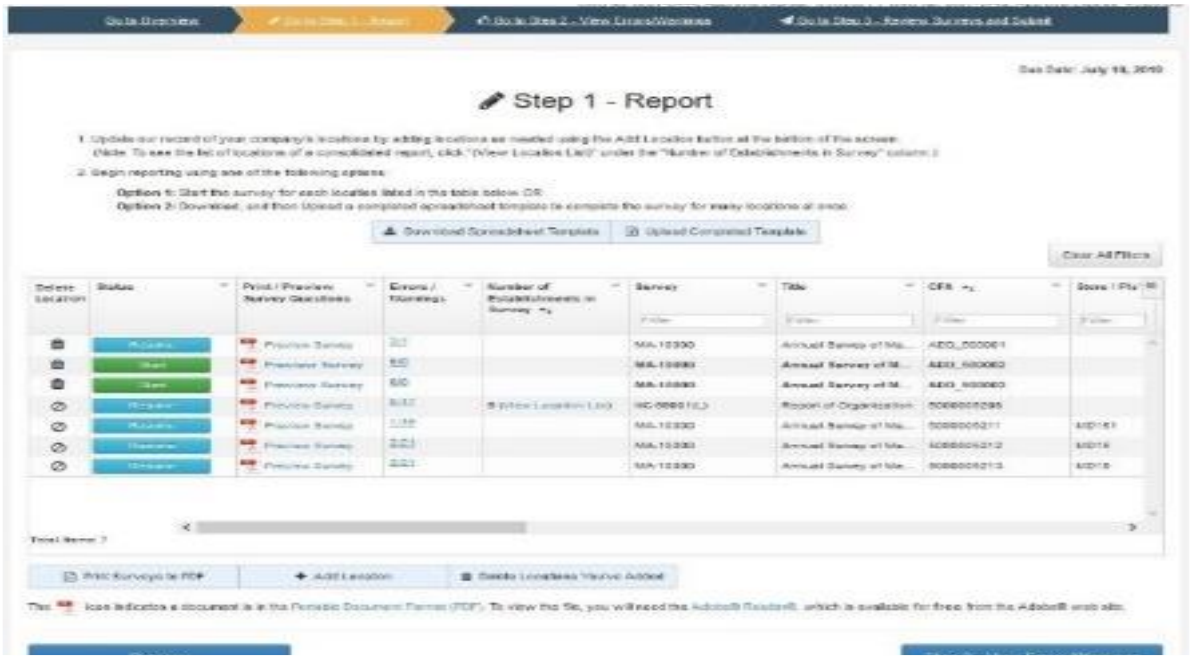


Figure 6. Proposed display for Deleting locations with trash can and prohibitory symbol design

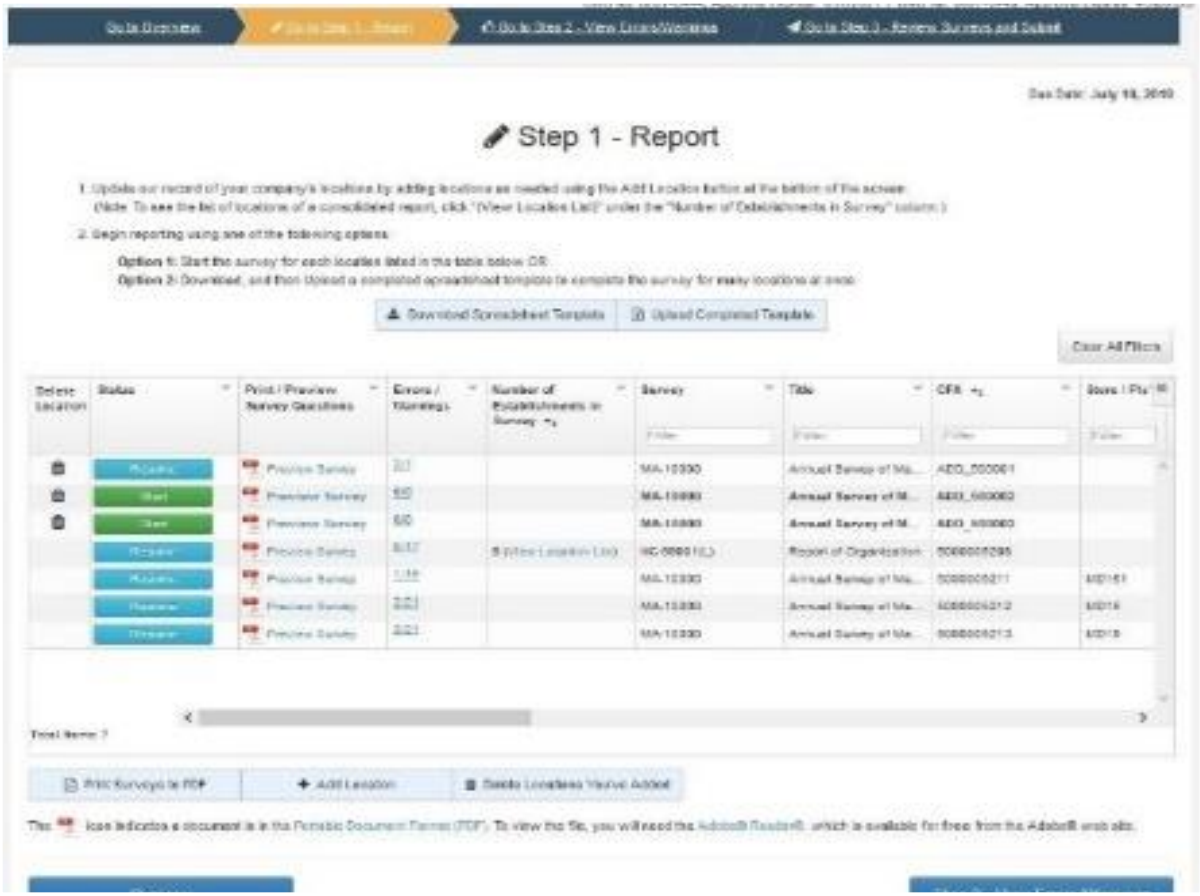


Figure 7. Proposed display for Deleting locations with trash can icon only design

Participants were successful in identifying the purpose of the trash can icon to delete locations. However, some were not clear that they were not deleting a pre-listed location, but a location they would have manually entered. This may be attributed to the low-fidelity testing environment in which we were evaluating prototype images of the design- participants did not actually engage in the process of manually adding a location to the listing.

Other symbols in the versions proposed for the redesign may not be universal to respondents. For example, participants commented that they were uncertain what the prohibitory symbol represented (see Figure 6). It may be the case that prohibitory symbols may indicate that respondents do not need to report to a location with this corresponding symbol, potentially causing them to bypass that report.

Additionally, during testing, participants readily identified the *Delete Locations You've Manually Added* button to remove locations from the listing- often before identifying the delete icon in the table- as outlined in the behavioral data shown in Figures 8-10 of versions tested. This indicates that the inclusion of the additional column for deleting locations may not be necessary as participants more readily noticed the alternative method for doing so on the screen.

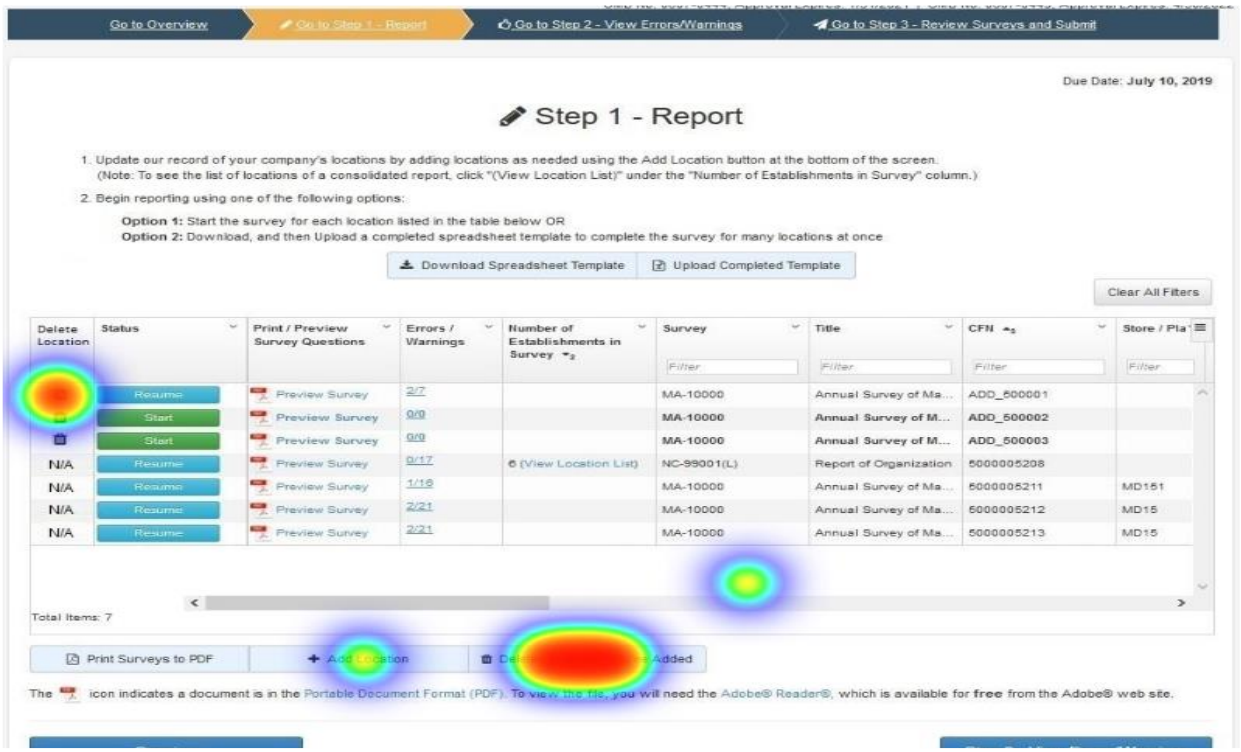


Figure 8. Behavioral data for Deleting locations with trash can icon and N/A design

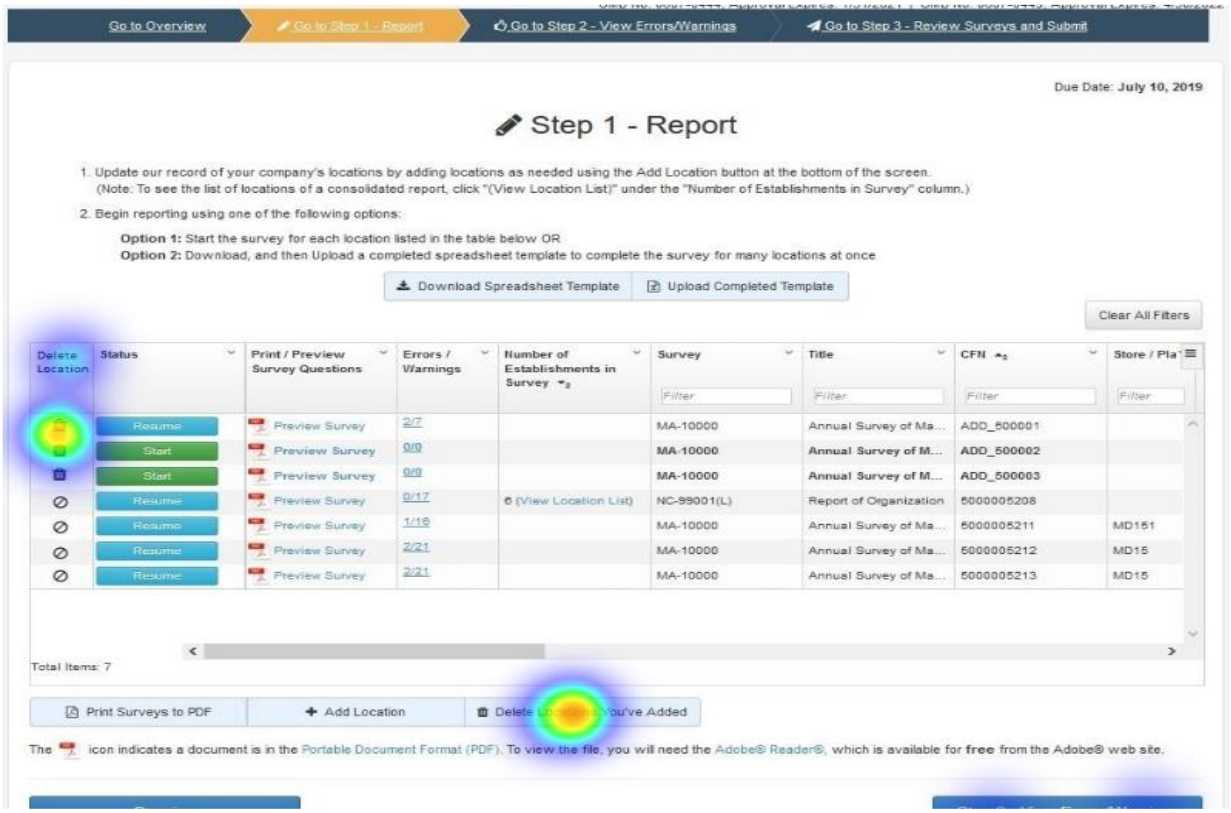


Figure 9. Behavioral data for Deleting locations with trash can and prohibitory symbol design

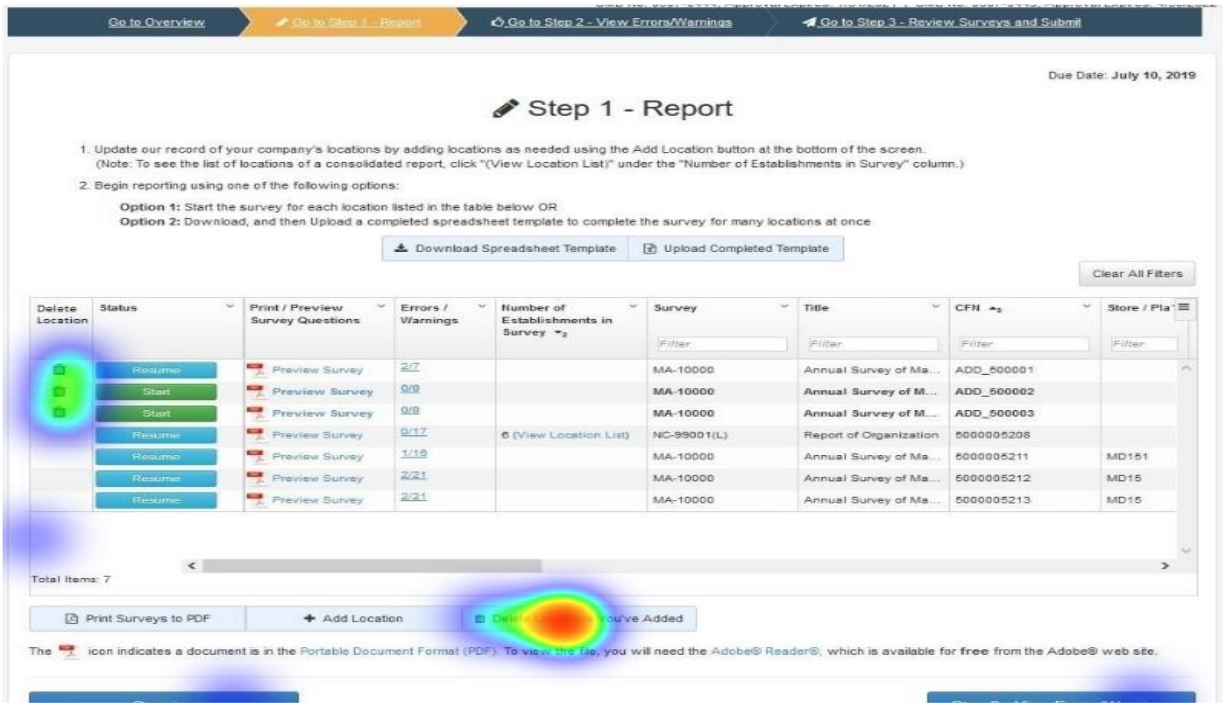


Figure 10. Behavioral data for Deleting locations with trash can icon only

Recommendation(s):

It may not be necessary to include an additional column for deleting locations. Adding another column to the table would shift current columns outlined in the table and may make some information less visible on the screen.

3.1.3. Confirming submissions

Task: You completed the requested survey and sent your data to the Census Bureau. You have logged back into the survey and are now back at the Report screen.

Identify which surveys have already been sent to the Census Bureau. How do you know that they have been sent? Click on any information that suggests your survey has been sent.

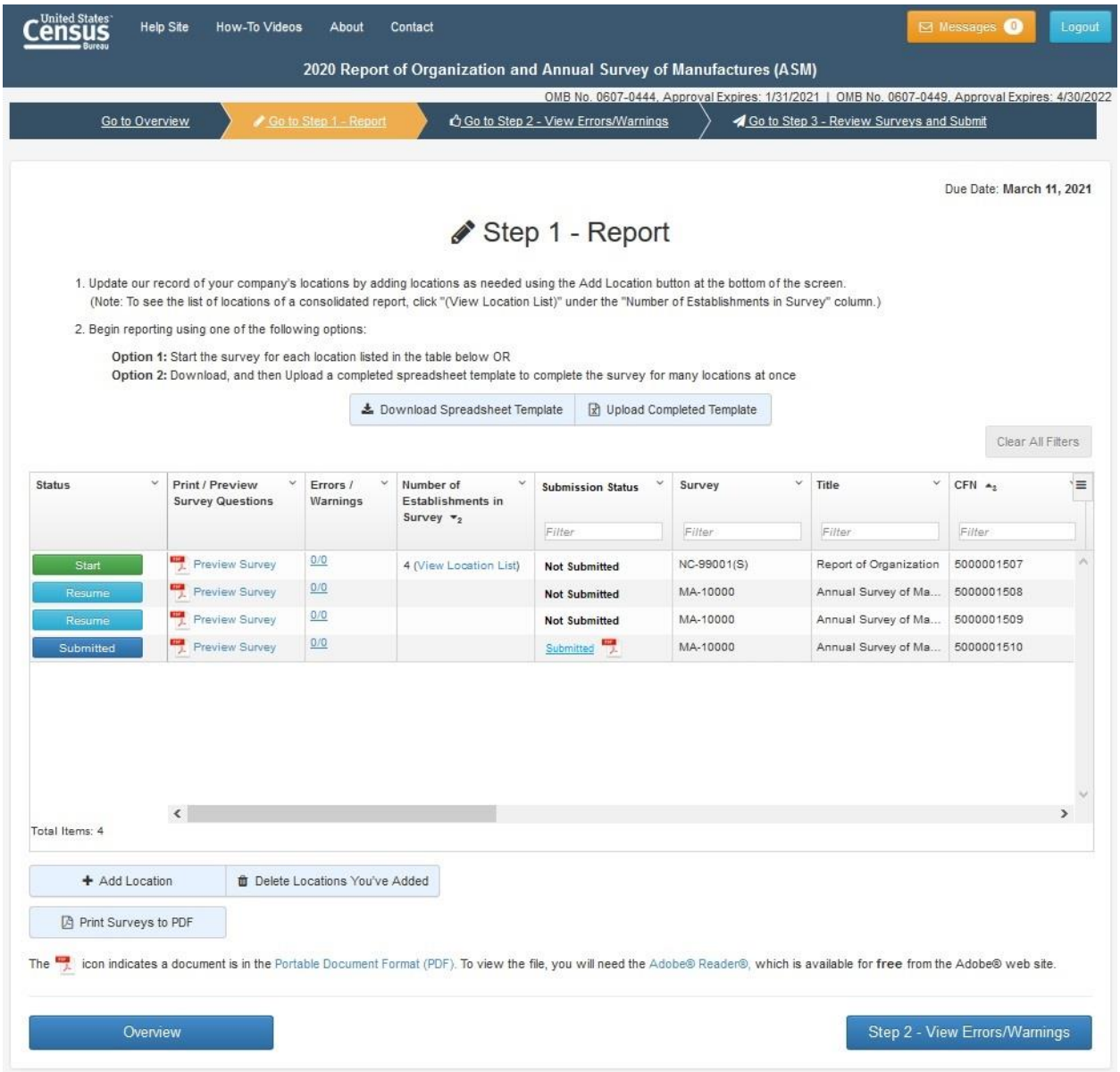


Figure 11. Proposed displays for submission details- Design 1

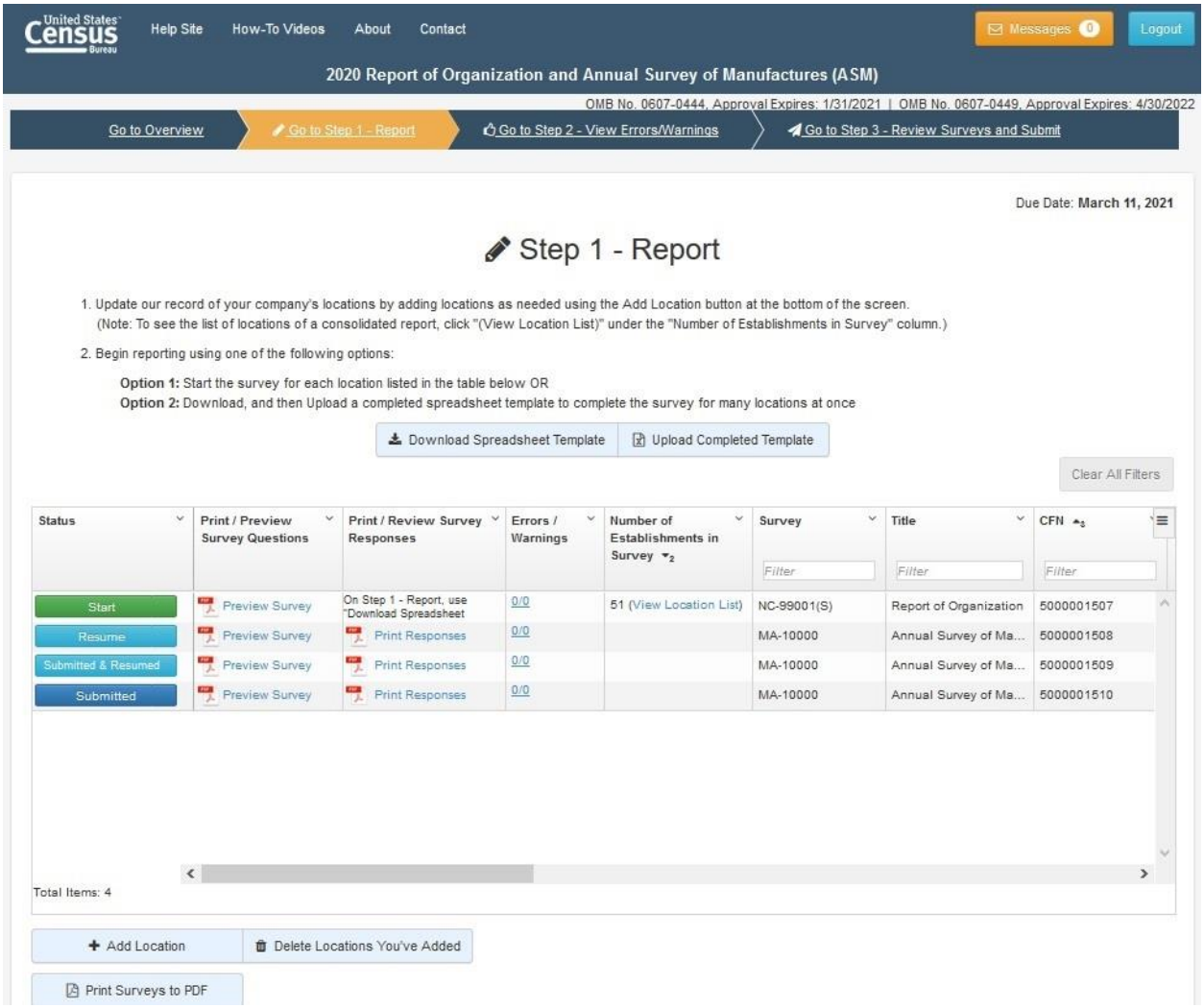


Figure 12. Proposed displays for submission details- Design 2

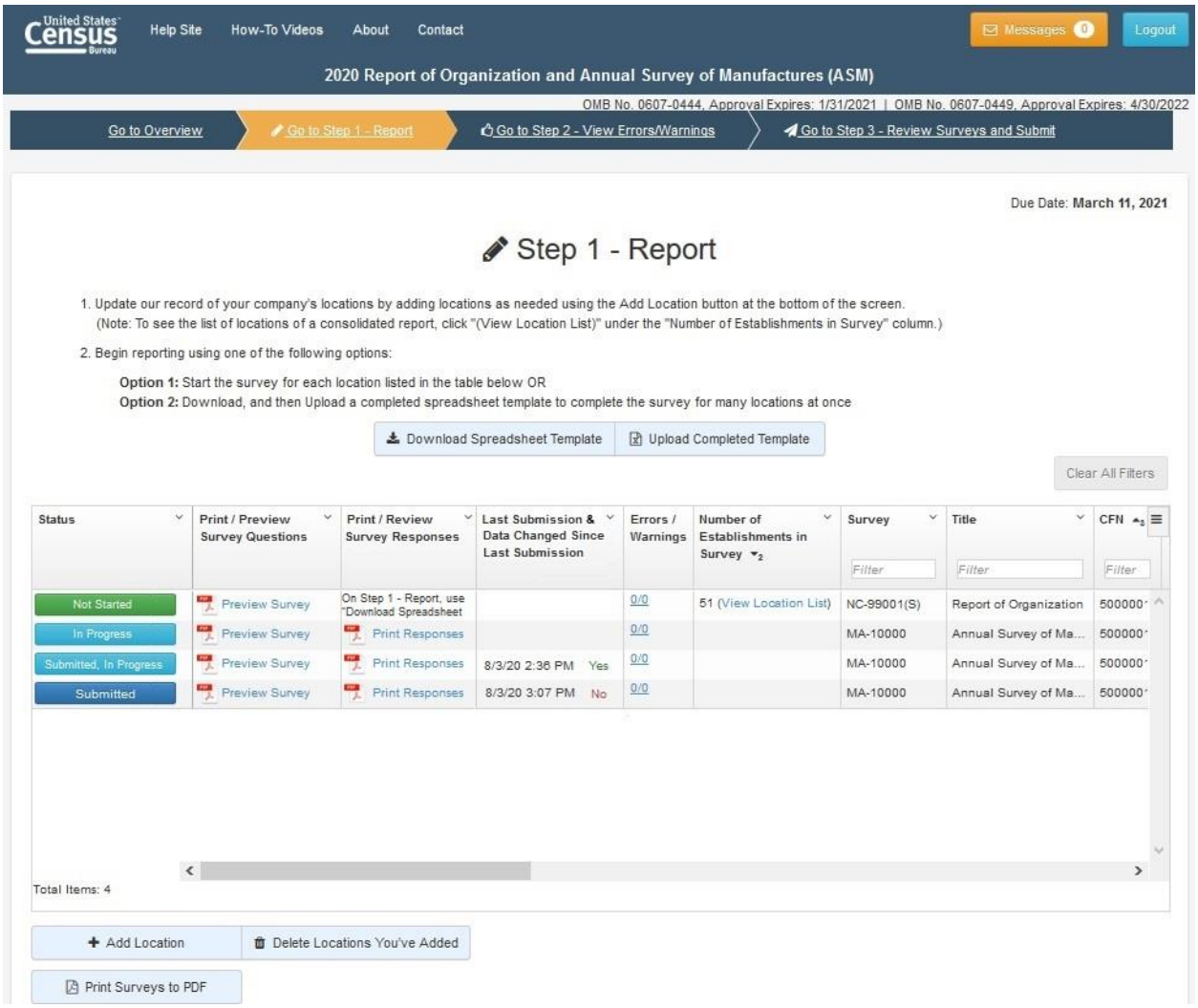


Figure 13. Proposed displays for submission details- Design 3

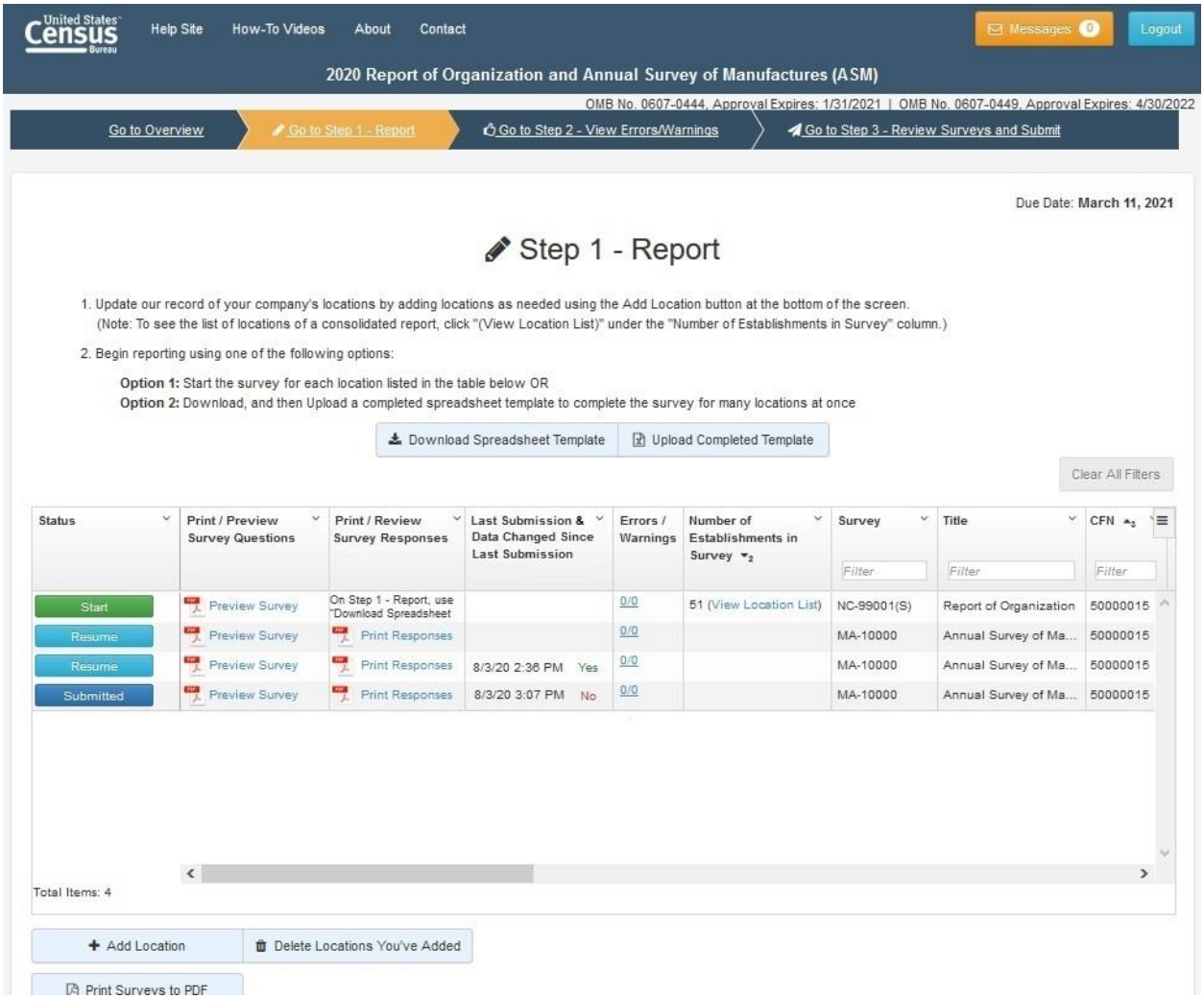


Figure 14. Proposed displays for submission details- Design 4

When asked to identify elements of the design that confirm a prior submission, most often, participants referenced the status button shown in the first column. The status buttons appear to be most noticeable, regardless of the version tested.

However, although most noticeable, the varying labels used for the buttons evaluated during testing were not as clear to participants. The most unclear labels identified were, *Submitted and Resumed* (Figure 12), *Submitted, Submitted; in progress* and *In progress* (Figure 13). Participants appeared to be less confident about their submission. It was not apparent that these labels were intended to inform them that they had gone back into an individual report without resubmitting, changing the overall submission status.

Additionally, links providing access to a downloadable copy of submitted responses in PDF format were not easily understood (e.g., *Submitted* in Figure 11, *Print Responses* Figures 12, 13 & 14). Participants did not understand that upon clicking on these links they would be provided a completed version of their report.

The Last submission and data changed since last submission shown in Figure 13 & Figure 14 were confusing as they contradicted the status buttons. For example, there is a last submission date although the status button states *Submitted, in progress.*

Recommendation(s):

Version 1 included in the evaluation may be the most easily understood by respondents. It has clear status buttons, and labels appear more distinct- Start, Resume, Submitted.

3.1.3.1. Total by EIN

When asked about their interpretation of the Total by EIN purpose, participants did not initially notice/reference this functionality on the screen. It became evident that the location of the tool was not apparent to participants. Some participants commented that they were initially looking at the lower portion of the screen near the table versus at the top.

When asked about the functionality of the tool, its purpose was clear from the description provided.

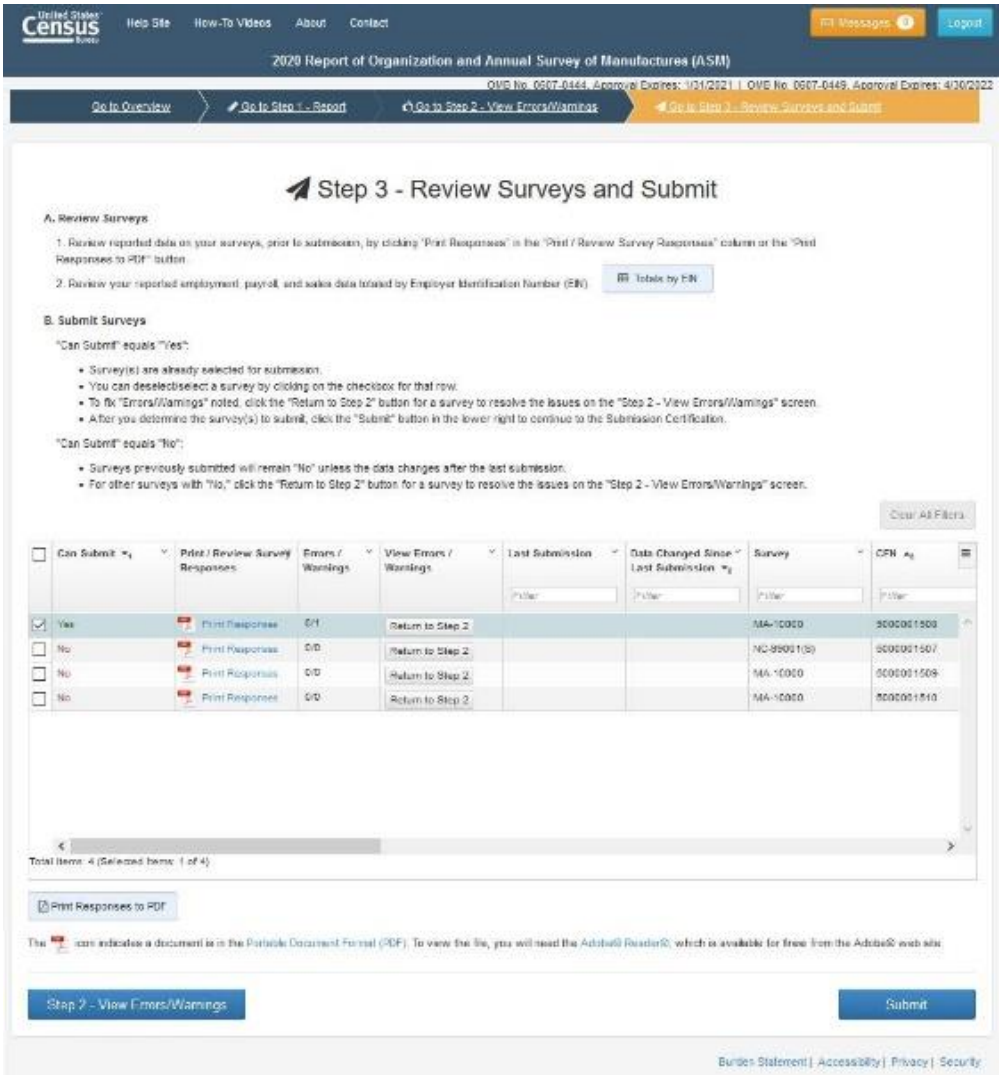


Figure 15. Total by EIN screen

Recommendation(s):

While the purpose of the *Total by EIN* functionality appears to be clear to participants, its placement on the screen makes it less noticeable. Consider moving the *Total by EIN* button to the lower portion of the screen by the table in closer proximity to where respondents are interacting with the interface, making it more noticeable and perhaps more likely to be utilized.

3.1.4. Pre-screening and NAICS

3.1.4.1. Overview screen

3.1.4.1.1. Content about Organization and Classification prescreening questions was missed from the Overview Screen

The Overview screen is intended to outline that respondents will need to complete a series of prescreening questions (inquiring about organization and classification) allowing the survey instrument to generate the most appropriate questionnaire for completion. However, outlining these details at the top of the screen as shown in Figure 16 was not apparent to participants. Participants did not realize that they would be asked prescreening questions- rather they reported that they would begin reporting- citing content under Step 1- Report. When asked about the content at the top of the screen referencing that they would be asked organization and classification questions on the upcoming screens, participants commented that they did not initially notice this content and they instead began reading under the Step 1 heading first.

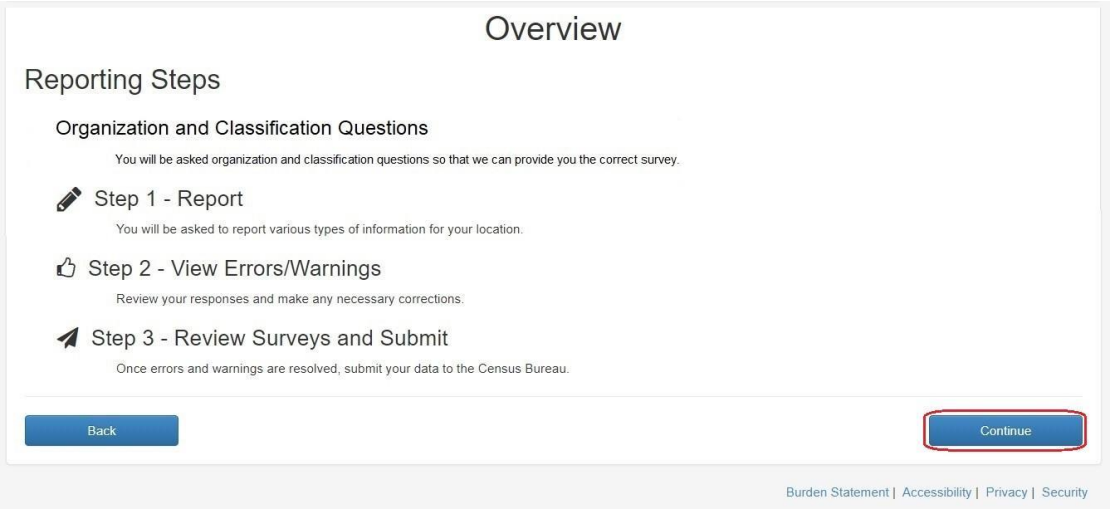


Figure 16. Round 1 Overview screen

Recommendation(s):

Move content outlining that respondents will be asked classification and organization questions under a Step sequence heading as these are where respondents may be more inclined to read first for instructions on getting started. Consider the use of short phrases, bullets, and words to make content appear less dense and readable.

3.1.4.2. NAICS: Generating a search (initial PBA selection screen)

3.1.4.2.1. Participants failed to enter details needed to generate a search

When identifying a Principal Business Activity (PBA) that was not pre-listed, participants appear to have understood that they needed to interact with the *Other* response category as shown in Figure 17. Most participants were able to select their sector from the drop-down menu and write in a response in the open field to describe their PBA. However, some participants failed to click on the radio button associated with the Other response category, instead they would immediately select the drop-down selection for sector. Additionally, there were participants who did select from the sector drop down listing but did not provide a write in response. They were able to move forward to the results screen without a prompt to enter these details.

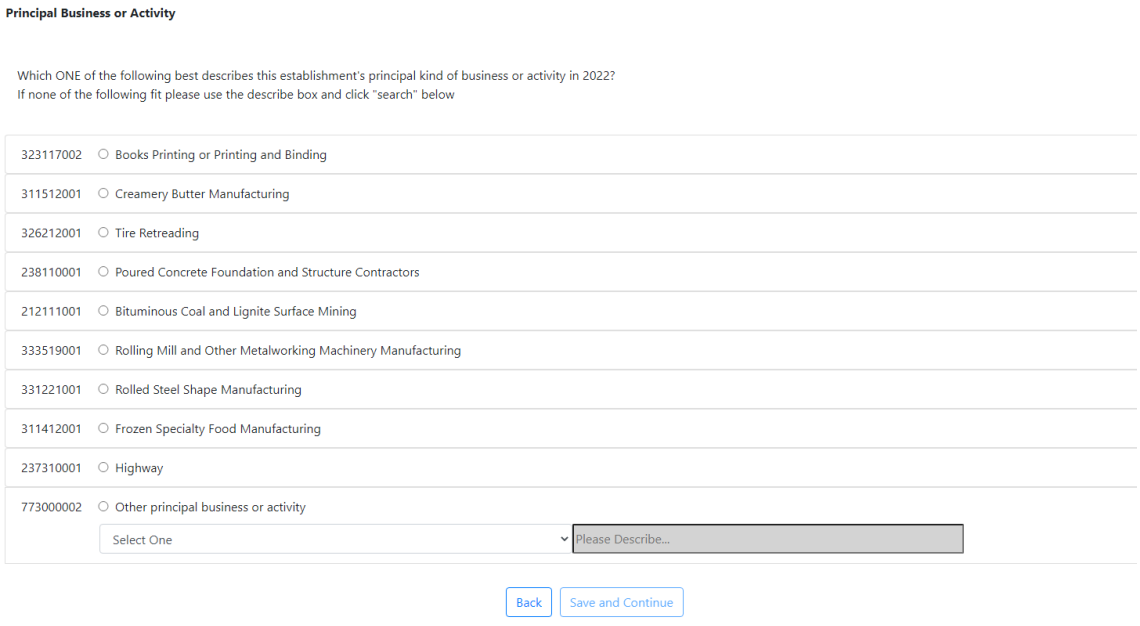


Figure 17. Round 1 PBA selection screen (initial)

Recommendation(s)

- Incorporate an edit message prompting a write in response if respondents fail to do so on the screen. Since the sector selection is not necessary for the machine learning functionality, this may not require an edit prompt.
- Include explicit instructions for interacting with the Other response category to indicate a PBA in the following locations:
    - Top of the screen with initial instructions
    - With-in write in field – gray instructional text

3.1.4.2.2. Radio button associated with the *Other* did not function as expected

Participants were unable to enable the *Other* response category by clicking on the associated radio button. The button would only become selected/enabled by selection from the drop-down menu or entering a write in in the open field. Often participants commented about the lack of this expected functionality and attempted to click on it several times before realizing it did not function in the manner in which they expected.

Recommendation(s):

Enable selection of the *Other* response category radio button upon clicking or interaction with the drop down or write in field.

3.1.4.2.3. Display altered based on browser settings

While screen sharing it was observed that the display of the screen varied across participants. For some participants the write in field shifted below the drop-down selection on the screen. This may have attributed to the findings cited in 3.1.4.2.1, where participants did not interact with both the write in field and drop-down menu as intended. With the write-in field placed directly below the drop-down menu, it may not have been seen as an independent field.

Recommendation(s):

Ensure that the design adapts to browser settings such that the display is not altered and remains consistent.

3.1.4.3. Machine learning results (6-digit NAICS)

3.1.4.3.1. Generating a new search query/modifying a write in was not immediately obvious

When asked how they would generate a new search from the 6-digit NAICS results screen, participants commented that they would go back to the initial PBA selection screen. They did not realize they could modify their text and generate a new search from the current 6-digit NAICS machine learning search results screen. This is likely attributed to the fact that when the screen required scrolling to view the complete listing of results, the search functionality is no longer visible on the screen.

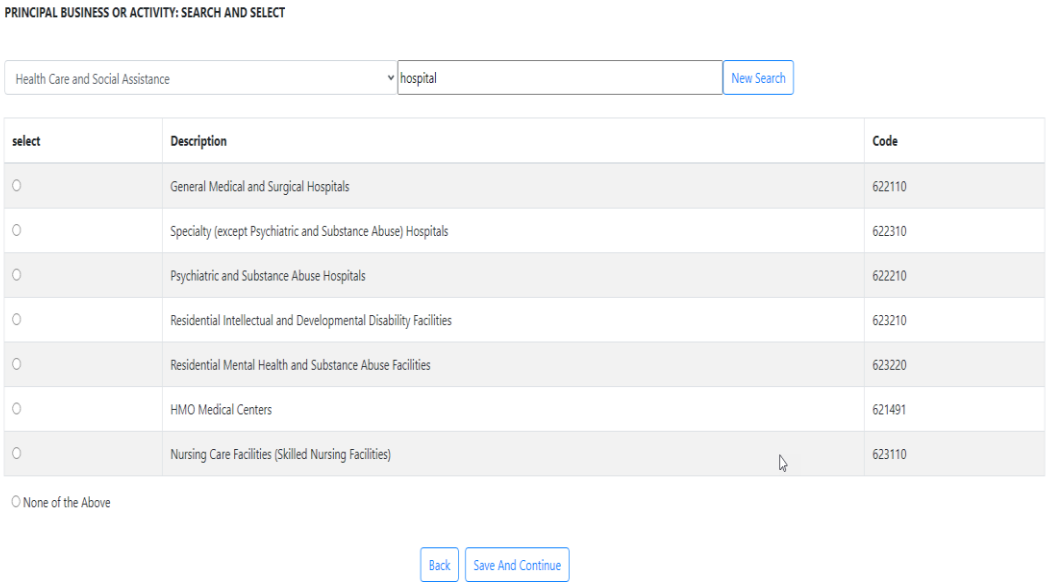


Figure 18. Machine learning results (6-digit NAICS). Search functionality is no longer visible when scrolling

Recommendation(s):

Maintain visibility of the search functionality when scrolling, making it stable on the screen.

3.1.4.4. Survey Loading

3.1.4.4.1. Expectation on how long it would take for a survey to be generated varied and the design lacked detail to help inform about expected duration

While participants did appear to understand that the purpose of this screen (Figure 19) was to indicate that their survey was being prepared, when asked how long they expected it to take- they were uncertain/provided varied responses.

Participants commented on the lack of information about the duration.

Figure 19. Round 1 Survey Loading screen

Recommendation(s):

- Indicate the time duration of how long a respondent should expect the download to take

Incorporate a visual indicator of progress- e.g., loading symbol as shown in Figure 20

Figure 20. Recommendation for Survey loading screen with loading symbol displayed

3.1.4.5. Survey Created

While participants understood that the purpose of the *Survey Created* screen (shown in Figure 21) was to indicate that their survey was ready, some were uncertain what would happen upon closing the screen.

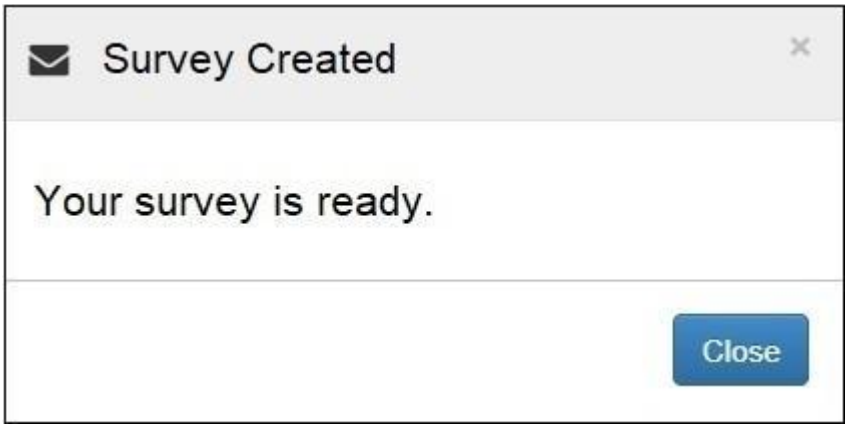


Figure 21. Survey Created screen

Recommendation(s):

Incorporate a modal design such that this screen appears as an overlay on the Report screen rather than a single stand-alone screen. Additionally, provide additional instruction/guidance on what respondents are to do or should expect upon closing the window. See Figure 22

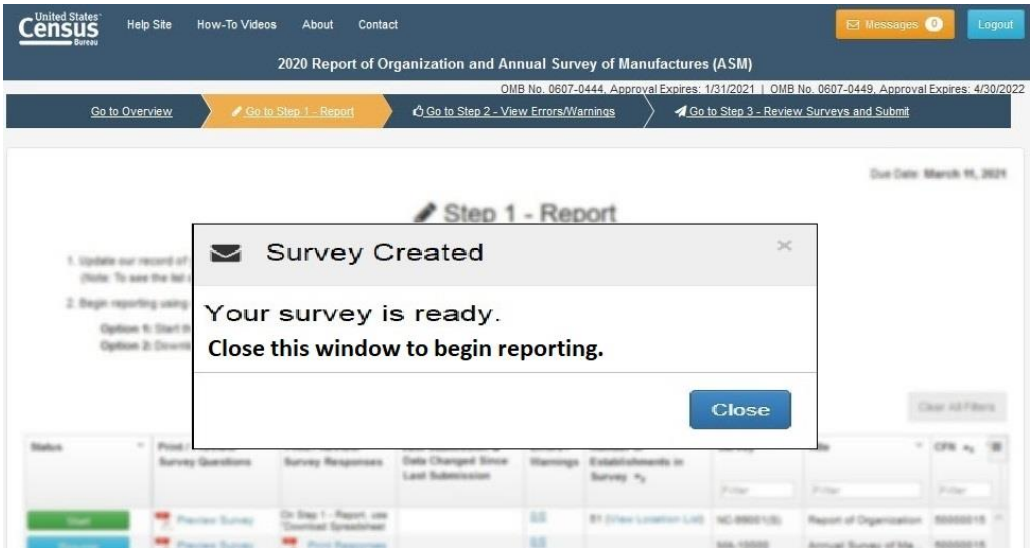


Figure 22. Recommended modal/overlay display for Survey creation screen

3.2. Round 2 Findings and Recommendations

3.2.1. Pre-screening and NAICS

3.2.1.1. Overview screen

3.2.1.1.1. A few participants failed to read the content about being asked organization/classification questions

Although most participants referenced that they would be asked organization/classification questions in the subsequent screens, some did not notice this content (as found in Round 1) although its placement was moved under the Step 1 heading. This may have been attributed to their prior experience with the survey impacting assumptions made about upcoming screens. These participants would often cite they would begin reporting information. Although not untrue, they failed to reference the types of information they will be asked to provide (e.g., organization, classification)

Recommendation(s):

Given respondents prior experience with the survey where they were not asked organization/classification questions, it may be helpful to visually alter the design of the Overview screen so that new content is more prominent. Consider the use of bullets to highlight new content, making it more prominent and less dense on the screen.

Additionally, it may be helpful to outline Step 1 as organization and classification- rather than report, making it clear that there are a series of questions for which the respondent has to respond prior to beginning their report as done in the past.

3.2.1.1.2. There was some uncertainty about what types of questions would be asked

As mentioned in 3.2.1.1.1, most participants did notice that they would be asked organization and classification questions in subsequent screens. However, some were unsure about what these questions may be.

Recommendation(s):

Include examples of the types of questions that may be asked in subsequent screens. This will be particularly helpful if such questions require the use of additional sources/records to complete.

3.2.1.2. NAICS: Generating a search (initial PBA selection screen)

3.2.1.2.1. Error messaging prompted write in responses when left blank

As outlined during round 1 testing, some participants failed to provide a write in response, making them unable to navigate to the next screen. In round 2 testing, an error message was implemented to help encourage write in responses. During round 2 testing, this feature assisted six respondents who initially left the write in field blank, prompting them to enter a description in the field.

Recommendation(s):

Maintain the error message. Consider if an error message prompt is needed for respondents who fail to select a sector from the drop-down selection.

3.2.1.2.2. Radio button not enabled upon selection

See Section 3.1.4.2.2

Note: This issue is likely to be resolved in a production environment

3.2.1.3. Machine learning results (6-digit NAICS)

3.2.1.3.1. Selection of “None of the above” may be a primary action when result selections are not applicable

When asked how they may proceed, if the results from the machine learning 6-digit NAICS results were not applicable, participants readily stated that they would select the *None of the Above* option. When asked if they noticed or would generate a new search, many commented that they would not be inclined to do so.

Recommendation(s)

None

3.2.1.3.1. Mapping onto a result selection was challenging for some participants

Participants had difficulty identifying a PBA from the available choices. There was a tendency to want to select multiple categories. Participants had trouble just mapping onto one category. Additionally, multi-unit establishments had to be reminded that they need only report for a given location, as they tend to want to report at a higher enterprise level, making the selection more challenging.

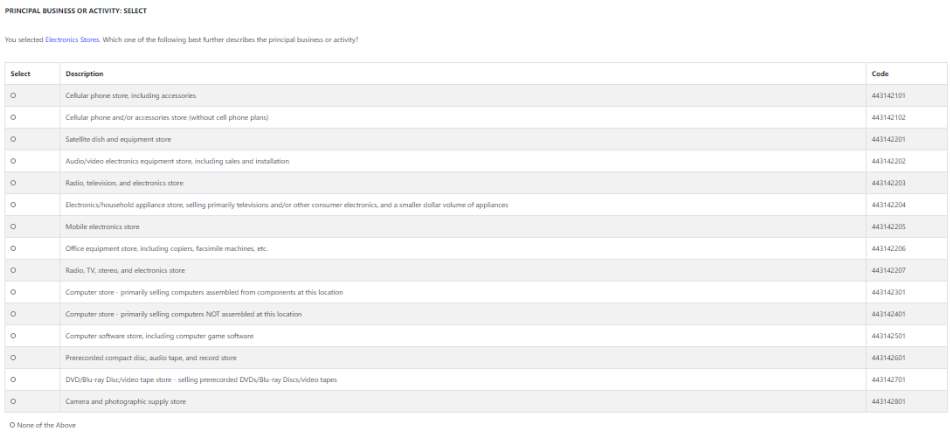
Recommendation(s)

Provide guidance for identifying one's PBA. Additionally, include the location information for MU's as a reminder of the establishment for which they are reporting. Respondents' will be asked to further refine their selection on the next screen.

3.2.1.4. 9-digit NAICS selection screen

3.2.1.4.1. Results displayed are repetitive for some, making its purpose unclear

The 9-digit NAICS selection (Figure 23) is needed to inform the appropriate survey a given respondent should complete by further refining their PBA classification. However, for some, the descriptions presented in the results were the same as those displayed from the prior 6-digit machine learning results query. For these participants, this was confusing. They often commented that this was "a confirmation" screen, or a "double check of what you've selected before" and did not understand the need to complete this additional screen. For others, who did receive a unique set of descriptors from which to select, they interpreted the 9-digit NAICS screen as a "drilling down" to a more descriptive activity, as intended.



PRINCIPAL BUSINESS OR ACTIVITY: SELECT

You selected Electronics Stores. Which one of the following best further describes the principal business or activity?

Select	Description	Code
O	Cellular phone store, including accessories	443142101
O	Cellular phone and/or accessories store (without cell phone plans)	443142102
O	Satellite dish and equipment store	443142201
O	Audio/video electronics equipment store, including sales and installation	443142202
O	Radio, television, and electronics store	443142203
O	Electronics/household appliance store, selling primarily televisions and/or other consumer electronics, and a smaller dollar volume of appliances	443142204
O	Mobile electronics store	443142205
O	Office equipment store, including copiers, facsimile machines, etc.	443142206
O	Radio, TV, stereo, and electronics store	443142207
O	Computer store - primarily selling computers assembled from components at this location	443142301
O	Computer store - primarily selling computers NOT assembled at this location	443142401
O	Computer software store, including computer game software	443142501
O	Prerecorded compact disc, audio tape, and record store	443142601
O	DVD/Blu-ray Disc/video tape store - selling prerecorded DVDs/Blu-ray Discs/video tapes	443142701
O	Camera and photographic supply store	443142801

O None of the Above

Figure 23. 9-digit NAICS/PBA selection screen

Recommendation(s):

Eliminate screen if 9-digit NAICS response selections are repetitive/not unique from the 6-digit NAICS description selected from the prior screen.

3.2.1.5. Survey loading

3.2.1.5.1. Loading symbol provided indicator of progress, but duration unknown

While the added loading symbol proposed after Round 1 testing (See Section 3.1.4.4) provided reassurance that they survey was being generated, participants varied in their expectation on the duration of time this may take. Some participants commented that having an indicator of time would be useful to have on the screen.

Recommendation(s):

Maintain the moving loading symbol as a visual indicator of survey creation progress. If the duration of survey creation is greater than 1-2 minutes, considering inserting text about the estimated time it may take to for this process to complete.

3.2.1.6. Survey Created

3.2.1.6.1. None

With the added modal design/screen overlay and instructions to *Close this window*, participants understood that closing the window allows them to begin reporting- interacting with the background screen. See Figure 22

4. Additional findings across rounds of testing:

- Participants were not initially aware that their write in response would be used to generate a search. It became evident once they reached the 6-digit NAICS machine learning search results. There was generally no negative feedback on this process. When asked about this functionality, some participants commented that it "made sense" and was similar to other experiences (e.g., tax software).

- Alternative search queries may be needed. Participants often commented about having the ability to search by a known NAICS code. Outlining this option on the initial PBA search may be helpful to respondents who know their NAICS code, improving their search query results. It is recommended that if searching by NAICS is an acceptable query, include instructions for this on the initial search.

- Most participants had relevant results from Machine Learning functionality. During round 2 testing at least 11 participants selected from top 5 results listed. Four participants experienced unsuccessful search queries in Round 1 testing; five participants experienced unsuccessful search queries in round 2.

- Some participants did not want to have to go through the process (selection of 6-digit and 9-digit NAICS) if they already mapped onto an existing NAICS. Instead, they wanted a reminder of existing NAICS from which to choose on the initial PBA selection screen.

- MU respondents were concerned about the prescreening and that they would have to go through this for each of their locations. Prescreening is not recommended for these establishments.

- If using a within-form view access to NAICS- there is uncertainty about what to expect after interacting with NAICS/PBA search (e.g., would respondents stay within the form, return to the Report screen?

- The purpose of the NAICS PBA selection process may not be clear within the instrument and needs to be effectively communicated so respondents understand its purpose. These details should be included on the PBA selection screen.

- There was some uncertainty about what survey/questions will be received if search functionality did not generate relevant/applicable results.

- Search queries were initially seen as irrelevant for some participants

- Round 1 examples:
  - Identified as Distributors
    - One participant, commented "Looking for one that mentions distribution"
    - Another participant's interaction:
      - First Search: Other→other→Write in: Distribution
      - Second Search: Wholesale trade→ Write in: Distribution→ Other misc. nondurable goods merchant wholesaler
- Round 2 examples:

Sector	Write-in description	Notes
Manufacturing	Manufacturing of glass for residential windows	6 and 9-digit NAICS results not applicable; ML 6 digit looking for description to mention residential windows
Manufacturing	1st search: Manufacture of doormats 2nd: manufacture textile doormats	Looking for something that references textile doormats;
Manufacturing	Automotive component	Wanted to select multiple categories- seating and electrical components
Real estate	Real estate investment rental leasing	Leased both residential and commercial- categories are distinct with 6-digit NAICS; but overarching category available with 9-digit NAICS results

Other services	1st search: Nonprofit 2nd search: private foundation	"none of these." Private foundation, do not solicit donation or fundraising, so none fit." 2nd search successful

4.1. SATISFACTION

The System Usability Scale (SUS) used to obtain satisfaction data is a simple, ten-item, 5-point Likert scale that provides a global view of subjective assessments of usability. Frequency results of the various element ratings are presented Figure 24 in and Figure 25, for round 1 and round 2 of testing, respectively.

#	Field	Strongly agree 1	2	3	4	Strongly disagree 5	Total
1	I think that I would like to use the site to complete the Economic Census	33.33% 9	33.33% 9	18.52% 5	11.11% 3	3.70% 1	27
2	I found the site to be simple	29.63% 8	33.33% 9	22.22% 6	7.41% 2	7.41% 2	27
3	I thought the site was easy to use	29.63% 8	37.04% 10	18.52% 5	7.41% 2	7.41% 2	27
4	I think that I could use the site without the support of a technical person	40.74% 11	37.04% 10	11.11% 3	3.70% 1	7.41% 2	27
5	I found that the various functions in the site were well put together	22.22% 6	44.44% 12	22.22% 6	7.41% 2	3.70% 1	27
6	I thought there was a lot of consistency in the site	33.33% 9	33.33% 9	22.22% 6	7.41% 2	3.70% 1	27
7	I would imagine that most people would learn to use the site very quickly	33.33% 9	25.93% 7	29.63% 8	7.41% 2	3.70% 1	27
8	I found the site to be very intuitive	18.52% 5	48.15% 13	22.22% 6	7.41% 2	3.70% 1	27
9	I felt very confident using the site	37.04% 10	33.33% 9	18.52% 5	7.41% 2	3.70% 1	27
10	I could use the site without having to learn anything new	48.15% 13	25.93% 7	11.11% 3	7.41% 2	7.41% 2	27

Showing rows 1 - 10 of 10

Figure 24. Round 1:Qualtrics output- SUS questions and frequency scores

#	Field	Strongly agree 1	2	3	4	Strongly disagree 5	Total
7	I would imagine that most people would learn to use the site very quickly	42.86% 9	33.33% 7	19.05% 4	0.00% 0	4.76% 1	21
3	I thought the site was easy to use	57.14% 12	33.33% 7	4.76% 1	0.00% 0	4.76% 1	21
6	I thought there was a lot of consistency in the site	60.00% 12	15.00% 3	20.00% 4	0.00% 0	5.00% 1	20
1	I think that I would like to use the site to complete the Economic Census	42.86% 9	33.33% 7	19.05% 4	0.00% 0	4.76% 1	21
4	I think that I could use the site without the support of a technical person	61.90% 13	28.57% 6	4.76% 1	0.00% 0	4.76% 1	21
8	I found the site to be very intuitive	52.38% 11	28.57% 6	14.29% 3	0.00% 0	4.76% 1	21
2	I found the site to be simple	57.14% 12	33.33% 7	4.76% 1	0.00% 0	4.76% 1	21
5	I found that the various functions in the site were well put together	42.86% 9	33.33% 7	19.05% 4	0.00% 0	4.76% 1	21
9	I felt very confident using the site	38.10% 8	47.62% 10	9.52% 2	0.00% 0	4.76% 1	21
10	I could use the site without having to learn anything new	47.62% 10	38.10% 8	9.52% 2	0.00% 0	4.76% 1	21

Showing rows 1 - 10 of 10

Figure 25. Round 2: Qualtrics output- SUS questions and frequency scores

APPENDIX A: Testing Protocol

Date _____

Time _____

Position title:_____

Tenure: _____

Company description:_____

Thank you for your time today. My name is XX and I work with the United States Census Bureau. My research team evaluates how easy or difficult Census products are to use. What works well, we keep. When potential users, such as you, have difficulty with something, we have an opportunity to fix it.

Thank you for signing the Consent form prior to your session. It explains the purpose of today's session and your rights as a participant. It also informs you that we would like to record the session to get an accurate record of your feedback. Only those of us connected with the project will review the recording and it will be used solely for research purposes. We plan to use your feedback to improve the design of an economic survey instrument and make sure it makes sense to respondents like you. If it is ok, I would like to start recording now.***[start CAMTASIA screen recording.]***

Thank you.

I am going to give you a little background about what we will be working on today. Today you will be helping us to evaluate the design of the online 2022 Economic Census instrument. The survey is in the early stages of development and we are interested in obtaining feedback to ensure that it has all of the information and features needed.

To do this, we will have you complete various tasks using the site. These will be consistent with tasks you would normally complete if you were requested to complete the actual 2022 Economic Census survey in the future. There are no right or wrong answers, we are mainly interested in your impressions both good and bad about your experience. I did not create the instrument so please feel free to share both positive and negative reactions.

The instrument is in its early design phase, so it is not interactive. Each screen will provide instructions for completing each task.

While you are completing the tasks I would like for you to think aloud. Essentially what this means is to verbally express what you are doing and experiencing as you use the site.

I may ask you additional questions about some of the screens you see today and your overall impressions.

Do you have any questions before we begin? Ok let's get started. I want you to read each task question aloud. Once you have completed the task just let me know by saying *finished* or *done*- then we can move on to the next task.

Task 1: Adding/Deleting locations (MU) (MU): Your company has gone through some changes and you need to make some updates to the location listing.

Part A: Your company has added 2 new locations. Please update your listing to reflect these additional locations.

- Success
- Fail

Task 2: Part B: The locations you added in Part A no longer exist. Remove the locations from the listing and Return to the Dashboard

- Success
- Fail

- MU: How easy or difficult was it to add locations? Why?
- Is it clear for which locations you are able to delete from the listing? Describe which can be removed/deleted.

Task 3: Any thoughts on available tools/features such as the Totals by EIN? What do you think is the purpose of this feature? Does the label make sense/do you have suggestions for changing the label of this feature?

Task 4: You've returned to the Report screen- identify which surveys have already been submitted.

- Success

- Fail

- How confident are you that your data was submitted successfully? Why?

- How easy or difficult was it to identify surveys that were submitted. Describe.

Prescreening and NAICS PBA Search

Option A: Prescreening

- *[Overview screen]* After reading this description, what do you expect to see on the next screen?

PBA Search

- This screen allows the respondent to select the PBA associated with their predetermined industry. If none of the options on the listing are applicable, how should one report? Can you report based on your establishment?

- After making a selection, the respondent clicks the Save and Continue Button. What do you expect to be displayed on the next screen?

- Is it clear what is being presented on this screen? Describe.

- If none of the PBAs are applicable, how should one proceed?

If OPTION A:

- Once your survey has been created you will receive the following message *[Survey created]*. Your survey will then appear on the Report screen
- What are your overall impressions of this functionality?

Round 2:

Option A: Prescreening

- **Overview screen:** After reading this description, what do you expect to see on the next screen?

- The next few screens will ask a series of questions about your establishment.

Please take a moment to review this screen. Is it clear what you are to do here? Describe. If none of the options on the listing are applicable, how should one report? Can you report based on your establishment?

- After making a selection, the respondent clicks the Save and Continue Button. What do you expect to be displayed on the next screen?

- Is it clear what is being presented on this screen? Describe.

- If none of the PBAs are applicable, how should one proceed?

Lets select an option from this screen and move forward.

9-digit NAICS screen:

- What are your first impressions of this screen. Is it clear what the purpose of this screen is? Describe.

- Is it clear what the difference is from the prior screen you saw?

- How would you proceed?

- What would you do if none of the results were applicable

What would you expect upon clicking to move to the next screen?

- Any thoughts? Is it clear what the purpose of this screen is? Describe.

- After answering the questions, you will be presented with the following screen. Any thoughts on this screen. Is it clear what is happening? What would be you expectation of how long this should take?
- Your survey will then appear on the Report screen
- What are your overall impressions of this functionality?

Option B: Within form view PBA

- Report screen

You want to begin reporting for a location via the form screen by screen view. You access the form by clicking on the Start button from the Report Now screen.

You encounter the following questions about your establishment.

- *Location*
- *Operational Status*

Next, you encounter a screen that ask about your Principal Business Activity. Please click on the link to access this screen.

Please take a moment to review this screen. Is it clear what you are to do here? Describe. If none of the options on the listing are applicable, how should one report? Can you report based on your establishment?

- After making a selection, the respondent clicks the Save and Continue Button. What do you expect to be displayed on the next screen?

- Is it clear what is being presented on this screen? Describe.

- If none of the PBAs are applicable, how should one proceed?

Lets select an option from this screen and move forward.

9-digit NAICS screen:

- What are your first impressions of this screen. Is it clear what the purpose of this screen is? Describe.

- Is it clear what the difference is from the prior screen you saw?

- How would you proceed?

- What would you do if none of the results were applicable

What would you expect upon clicking to move to the next screen?

- Any thoughts? Is it clear what the purpose of this screen is? Describe.

- After answering the questions, you will be presented with the following screen *[Survey loading]*. Any thoughts on this screen. Is it clear what is happening? What would be you expectation of how long this should take?

What do you expect to happen next when you go to the next screen in the form?

General:

[Review Satisfaction responses- probe about any items rated <3)]

- Overall, what was your impression of the site? Is there anything you liked / disliked about the site?

- In general, would you say that it was easy or difficult to complete the tasks using the site? Which tasks, if any, were difficult to complete?

- If you could change anything about the site, what would that be?

- If you were having trouble with using this site, would you be interested in using a chat support feature, or would you prefer another way to get help? *If yes to chat support,* describe how you envision this working? Any potential challenges you see with something like this?

If time permits:

- (MU/SU) Based on the available reporting options, how has your establishment reported in the past? Why? How was your experience reporting this way?
- *[MU]If reported via spreadsheet in the past* (MU) How would you describe the process of downloading and uploading a spreadsheet?
    - How may we improve your experience with the spreadsheet functionality?
    - (MU/SU) *If reported via form view in the past*: What are your thoughts on using the form to report to the survey?

- Review the three options for going between screens in the form. Is it clear for which screens you are able to visit? Describe which screens you are able to revisit.
- Which of the displays would you prefer? Why?

- Is there anything else you would like to mention that we haven't talked about?

This concludes our session. Thank you for your time and valuable feedback.

APPENDIX B: User tasks for Round 1 testing

Task [Adding locations]: Your company has gone through some changes and you need to make some updates to the location listing.Your company has new locations. You need to update your listing to reflect these new locations that aren't listed.
Click on where you would go to do this.

Task [Deleting locations]: The locations you added in the prior question no longer exist. You want to remove the locations from the listing. Is it clear which location(s) you are able to remove. Click on where you would go to remove them.

Task [Submission details]: You completed the requested survey and sent your data to the Census Bureau. You have logged back into the survey and are now back at the Report screen.

Identify which surveys have already been sent to the Census Bureau. How do you know that they have been sent? Click on any information that suggests your survey has been sent.