

August 29, 2006

2005 National Census Test: Bilingual Form Analysis

Julie A. Bouffard
Jennifer Guarino Tancreto

Decennial Statistical Studies Division

CONTENTS

EXECUTIVE SUMMARY	iii
1. INTRODUCTION	1
2. METHODOLOGY	2
2.1 Panel Design	2
2.2 Response Mode	2
2.3 Mailing Strategy	3
2.4 Sample Design and Standard Errors	3
2.5 Analytical Methods	4
2.6 Calculation of Self-Response Rates	4
2.7 Calculation of Item Nonresponse Rates	5
2.8 Quality Assurance Procedures	6
3. LIMITATIONS	6
3.1 Population of Inference	6
3.2 Confounding Factors	6
3.3 Census Test Environment	7
3.4 Natural Disasters	7
4. RESULTS	7
4.1 Self-Response Rate Comparisons	7
4.2 Bilingual Form Response Patterns	9
4.3 Item Nonresponse Rate Analysis	10
4.4 Comparisons of Demographic Characteristics	17
4.5 Public Reaction to the Bilingual Form	19
5. CONCLUSIONS AND FUTURE CONSIDERATIONS	19
ACKNOWLEDGEMENTS	21
REFERENCES	22

LIST OF TABLES

Table 1. Self-Response Rates and Standard Errors (in percent) by Language and Response Mode at the National Level and by Strata.....	8
Table 2. Item Nonresponse Rates and Standard Errors (in percent) by Language Treatment.....	11
Table 3. Item Nonresponse Rates and Standard Errors (in percent) by Language Treatment and Strata	13
Table 4. Item Nonresponse Rates and Standard Errors (in percents) by Language Treatment and Person 1 Hispanic/Non-Hispanic	14
Table 5. Item Nonresponse Rates and Standard Errors (in percent) by Language Column on the Bilingual Form.....	15
Table 6. Item Nonresponse Rates and Standard Errors (in percent) by Language Column for Person 1 Hispanic on the Bilingual Form	16
Table 7. Demographic Characteristics and Standard Errors (in percents) by Language Treatment...	17
Table 8. Demographic Characteristics and Standard Errors (in percents) by Language Column on the Bilingual Form	18

EXECUTIVE SUMMARY

In late 2005, a mailout/mailback national test was conducted using variations of questionnaire content and various methods to increase response to the Census. One of the experimental treatments in the 2005 National Census Test was a bilingual questionnaire with a “swim lane” design. The swim lane design provided two response columns, one in English and one in Spanish, each containing the same questions and response categories. This form was mailed to a randomly selected set of 10,000 housing units.

The desire to test a bilingual form stemmed from the hope that this form would increase response to the census as well as lower item nonresponse by meeting the needs of a broader audience than the standard English form. The 2010 Language Research and Development Group was also interested in any public reaction to the bilingual form, specifically any backlash in areas that were predominantly non-Hispanic.

Results

Self-Response Rates

- Both experimental panels offered households the option of responding by Internet or a paper questionnaire. There were, however, differences in the administration of the Internet between the bilingual and English panels. Most notably, there was no Spanish Internet instrument in the 2005 test, and the placement of the invitation to use the Internet varied across the panels. We cannot conclusively determine the impact of these differences on response behavior. Therefore, we present two distinct measures for response rates comparisons, paper response and total response.
- We first present paper response rates since that paper form contained the experimental bilingual design, and the Internet did not. The bilingual panel achieved significantly higher paper response rates (by 2.2 percentage points) compared to the English panel.
- The paper response rate increase was even more evident in areas with a high concentration of the non-White or Hispanic populations, with a 3.2 percentage point increase. We also found that the bilingual form showed a significant increase (1.7 percentage points) for paper response in areas with a low concentration of the non-White or Hispanic populations relative to the English form.
- Secondly, we studied total response (paper and Internet) to the panels. The bilingual form also received significantly higher total response compared to the English only panel, both nationally (by 1.1 percentage points) and for areas with a high concentration of the non-White or Hispanic populations (by 2.8 percentage points). The bilingual form had no impact on total response in areas with a low concentration of non-White or Hispanic populations.

Bilingual Form Use

- Of the respondents who filled out the bilingual paper form, 2.7 percent predominantly used the Spanish language column to complete the form. This figure jumped to 8.3 percent in areas with a high concentration of the Non-white or Hispanic populations.
- Less than 1 percent of bilingual form respondents used both language columns to respond. Thus, it appears that the vast majority of respondents were able to use the bilingual form as intended.

Item Nonresponse Rates by Panel

- Item nonresponse rates for the bilingual panel were significantly higher than the English only panel for household count, tenure, undercount, telephone number and Hispanic origin. No item nonresponse rate differences were found for other person items that were compared in this analysis.
- We observed the same results for both strata, except that the race item also had significantly higher item nonresponse for the bilingual panel in the high non-White and Hispanic concentration stratum.
- The bilingual panel had a significantly higher percent of forms with all household level items missing (2.1 percent) compared to the English panel (0.4 percent). This finding may suggest a forms design issue with the bilingual form, in that respondents of the bilingual form may have missed the first page of the questionnaire (where the household items are located).

Item Nonresponse Rates – Bilingual Form Only

- Looking only at the bilingual form, we found that item nonresponse rates for household count, tenure, undercount, race and overcount were significantly higher for the Spanish column as compared to the English column.
- Respondents that used the Spanish column had a significantly higher proportion of returns with all household items missing (5.9 percent) compared to respondents that used the English column (2.0 percent). Again, there may be a form design issue with the bilingual form driving these results.

Comparison of Demographic Characteristics

- We observed a few demographic differences between English form returns and bilingual form returns. Household size and the proportion of Hispanics were both significantly higher for the bilingual form. The proportion of males was significantly lower in the bilingual panel compared to the English only panel.
- We then compared demographics across the two language columns on the bilingual form. A significantly higher proportion of persons who used the Spanish column were renters, male,

Hispanic, younger, of some other race and from larger households than those who used the English column.

Public Reaction

- There was no formal evaluation to scientifically measure public reaction to the bilingual form. Instead, the Decennial Management Division recorded any questions or comments from the public that pertained to the use of the bilingual form. They noted less than ten inquiries about the bilingual questionnaire. Almost all of these respondents inquired as to why they received an English and Spanish questionnaire.
- Data also showed no sign of backlash, in that, we did not see any drop in self-response rates, especially in areas with a low concentration of non-White or Hispanic populations. We cannot say, however, what reaction would occur if the bilingual form were distributed to a larger audience within the context of a decennial census.

1. INTRODUCTION

In preparation for the 2010 Census, the Census Bureau is conducting a series of tests. In late 2005, a mailout/mailback national test was conducted using variations of questionnaire content, and various methods to increase response to the Census, including replacement questionnaire methods. The test also included the Internet as an optional mode for completing the census short form. Census Day was September 15, 2005.

The objectives for the 2005 National Census Test were:

- Test methods to improve completeness and accuracy of reporting for short form items, including tenure, relationship, age and date of birth, and race and Hispanic origin,
- Test ways to reduce respondent and data capture errors, and improve respondent friendliness in mail and Internet modes,
- Test ways to improve coverage accuracy by reducing omissions and erroneous enumerations, and/or flagging potential errors for coverage followup interviews,
- Test ways to improve the operational feasibility of the second mailing,
- Test ways to improve self-response and maintain data quality by mailing bilingual questionnaires.

As mentioned above, one of the experimental treatments in the 2005 NCT was a bilingual questionnaire with a “swim lane” design. The swim lane design provided two response columns, one in English and one in Spanish, each containing the same questions and response categories (see Appendix A for pages from the bilingual form). The bilingual form instructed respondents to choose the language that was most comfortable for them. This form was mailed to a randomly selected set of housing units that formed an experimental panel.

The desire to test a bilingual form stemmed from the hope that this form would increase response to the census as well as lower item nonresponse by meeting the needs of a broader audience than the standard English form. Another hope was that the use of a bilingual form would reduce the workload associated with fulfilling requests for Spanish forms during the census. Note that we were unable to measure any potential reduction in this test since requests for language forms were not fulfilled.

Specifically, the 2010 Language Research and Development (R&D) Group at the Census Bureau proposed this research to answer the following research question: *What is the impact of a bilingual questionnaire on response behavior for a national sample of housing units?* In addition, the 2010 Language R&D Group was interested in any public reaction to the bilingual form, specifically any backlash in areas that were predominantly non-Hispanic.

The 2005 NCT was not the first time a bilingual form was tested in the context of a mid-cycle census test. The Spanish Forms Availability Test (SFAT) took place as a part of the 1993 National Census Test and consisted of three panels: a control panel with an English only questionnaire, a panel that distributed two distinct questionnaires (one in English and one in Spanish) in the same mailing package, and a panel offering a bilingual form with a back-to-back design (de la Puente *et al.*, 1994a). The SFAT also included a telephone debriefing with

respondents who had received both the English and Spanish forms to gauge their reactions (de la Puente *et al.*, 1995).

Results from the SFAT suggested that offering a Spanish questionnaire significantly increased response to the census test by roughly four to five percentage points in areas where more than 30 percent of households were known to be linguistically isolated¹; however, the results also showed that Hispanics were more likely to omit information on the Spanish form (bilingual or separate form) compared to Hispanics who responded to the English form (Corteville, 1994; de la Puente *et al.*, 1994a).

The bilingual experiment in the 2005 NCT extends this research by testing a new bilingual form with a swim lane design (i.e., two column design) instead of a back-to-back design.

2. METHODOLOGY

2.1 Panel Design

The 2005 NCT was comprised of 20 experimental panels². We used two of these panels to assess the effectiveness of the bilingual form. We used the bilingual form panel of 10,000 sampled housing units as the experimental panel, and an English form panel of 30,000 sampled housing units that was most similar in content to the bilingual questionnaire as the control panel. The English panel used as a control in this analysis also contained experimental treatments for the Hispanic origin, race, tenure, and age items. The bilingual form contained the same experimental questions for Hispanic origin and race as the control, but did not include the experimental version of the tenure or age questions. See Appendices A and B for pages from the forms.

2.2 Response Mode

Housing units in each 2005 NCT experimental panel, including the bilingual form panel, were invited to respond using the Internet. The Internet form did not, however, contain all of the experimental treatments and, most notably, did not provide the option of responding in Spanish³. In this sense, households that responded by the Internet were no longer considered part of their original experimental panel since they were not exposed to the experimental treatment(s) embedded in the paper questionnaire. Therefore, most analyses focus on responses returned via the paper questionnaires, and exclude any households that responded by the Internet.

¹ A linguistically isolated household was a household where no household member 14 years of age or older spoke English or spoke English very well (de la Puente *et al.*, 1994b).

² For more detail on the panel design refer to U.S. Census Bureau (2005). Note, the bilingual panel corresponds to Panel 20 and the English only panel corresponds to Panel 4.

³ In the English and Spanish columns of the bilingual form, the Internet invitation contained a disclaimer that the web site was "English only."

2.3 Mailing Strategy

The 2005 NCT used multiple mailings to contact sampled housing units. Every housing unit was sent an advance letter as a first contact. The advance letter informed households that they would soon receive a request to complete a questionnaire for the 2005 National Census Test.

The second mailing was an initial questionnaire package. Housing units received a paper questionnaire and a first-class postage-paid return envelope. Also included in the mailing package was a letter from the Census Bureau's Director that encouraged households to respond and provided the option of responding by Internet. Households selected for the bilingual panel received the bilingual questionnaire, and households selected for the English panel received the English form. The English questionnaire reminded respondents of the option to respond by the Internet before the first question, but the bilingual questionnaire did not.

The third mailing was a reminder postcard. The reminder postcard included a statement reminding households to respond to the census test if they had not already done so. It also provided instructions so that households could use the Internet to respond. Presenting the reminder postcard content in English and Spanish in the bilingual swim lane format required more space than was available on the postcard; therefore, the reminder postcard mailing was actually a letter for the bilingual panel.

The fourth and final mailing was a targeted replacement questionnaire. A replacement questionnaire that looked identical to the initial questionnaire (i.e., contained the same experimental treatments) was sent to all housing units that had not responded prior to September 13, 2005. Accompanying the questionnaire was a letter from the Director urging response and providing instructions for using the Internet.

Note that there was no telephone or personal visit followup for nonresponding households in the 2005 NCT.

2.4 Sample Design and Standard Errors

The housing units in the 2005 NCT were selected from mailout/mailback areas of the country. The housing units selected for the bilingual form panel (10,000 housing units) and its corresponding control panel (30,000 housing units) were equally allocated to two strata that reflect differences in the racial and ethnic composition, and, hence, response propensity of the mailout/mailback universe. The high non-White or Hispanic concentration stratum, which encompassed roughly 32 percent of housing units in the universe, contained a high proportion of the non-White populations (i.e., Black, Asian, etc.) or Hispanic populations. The remaining 68 percent of the housing units fell in the low non-White or Hispanic concentration stratum. For more information about the creation of the strata, please refer to memorandum by Bentley (2005). All estimates in this report are weighted to account for the oversampling of the high non-White or Hispanic stratum.

We computed standard errors for all estimates using a stratified jackknife replication procedure. This computation method accounted for the stratification in the sample, which we expect to lower the standard errors compared to a simple random sample. Clusters of housing units (or

housing units selected at each hit) were assigned sequentially to one of 250 replicates. This assignment approach also accounted for the clustering of persons within a household in computing errors for person level estimates, since persons within households were contained in the same replicate.

2.5 Analytical Methods

We used the following comparisons to evaluate the effectiveness of the bilingual form:

- Self-response rate comparisons (Section 4.1)
- Bilingual form response patterns (Section 4.2)
- Item nonresponse analysis (Section 4.3)
- Comparisons of demographic characteristics (Section 4.4).

We tested for differences in the self-response rates, item nonresponse rates, and demographic characteristics between the bilingual panel and the English panel. Computed differences were compared to critical values. If the computed difference was greater than or equal to the critical value, then the difference was deemed significant.

This report also includes a qualitative summary of public reaction to the bilingual form (Section 4.5).

2.6 Calculation of Self-Response Rates

The self-response rate is a measure of respondent behavior with regard to responding to the census test.

The denominator is the number of sample housing units minus those cases identified by the United States Postal Service (USPS) as “undeliverable as addressed” (UAA). Cases that were UAA⁴ were defined as those housing units where there was no response (paper or Internet), and both the initial questionnaire and replacement questionnaire mailings were flagged as UAA. Any housing units determined to be UAA were considered ineligible units. We observed an 8.1 percent UAA rate for the bilingual panel, and a 7.7 percent UAA rate for the English panel. As expected, these rates were not significantly different from each other.

The numerator is the number of sample households for which we received a nonblank return. A census return was denoted as “blank” if fewer than two⁵ of the following items were completed: tenure, household count, name, relationship, sex, age or date of birth, Hispanic origin, race, and ancestry.

Also, we limited the numerator to primary returns. We selected a primary return when multiple responses were received for a given housing unit, using the following rules⁶:

⁴ See memorandum from Rothhaas (2004a).

⁵ The item completeness criteria are listed in the memorandum from Rothhaas (2004b).

⁶ For more information about the selection of primary returns, please see the memorandum from Rothhaas (2004c).

1. When more than one paper return was received from a single household (i.e., we received both an initial and replacement questionnaire return), we selected the first nonblank form received. In the rare case that two nonblank paper forms were checked in on the same date and in the same batch, we selected the initial questionnaire return.
2. When multiple Internet returns were received from a single household, we selected the first nonblank return.
3. When paper and Internet returns were received for a single household, we selected the first nonblank return based on date received (i.e., check-in date/submitted date). If a nonblank paper and nonblank Internet return were received on the same day from the same household, we selected the paper return as it was probably mailed before the Internet was submitted.

The formula for the self-response rate is presented below.

$$\text{Self-response rate} = \frac{\# \text{ of nonblank, primary returns}}{\text{panel sample size} - \text{UAA for the panel}} * 100\%$$

Please note that the self-response rate presented above corresponds to the rates used in previous census tests, including the 2003 National Census Test, the Census 2000 experiments, and the 1992 and 1993 Census Tests. We used the self-response rate because it is not subject to variation in UAA rates. Specifically, the denominator of the self-response rate excludes cases for which eligibility cannot be determined, such as units that are UAA. Therefore, any variation in the UAA rates across panels did not contribute to differences in the self-response rates.

Lastly, please note that the self-response rate defined here is not directly comparable to the Census 2000 mail response or mail return rates. The self-response rate is not a return rate in the sense that we do not definitively know the occupancy status of housing units included in the denominator or the status of cases that are excluded as UAAs.

2.7 Calculation of Item Nonresponse Rates

Item nonresponse rates were computed as indicators of potential data quality issues. The analysis of item nonresponse rates was restricted to nonblank, primary paper returns for this analysis. The item nonresponse rates were calculated according to the following definition:

$$\text{Item nonresponse rate} = \frac{\# \text{ of records with missing data for a particular item}}{\text{total number of records}} * 100\%$$

Item nonresponse rates were calculated for both housing unit level items and person level items. For housing unit level items, the term “records” refers to housing units. The total number of housing units was defined as the number of housing units from all nonblank, primary paper returns. For person level items, the term “records” refers to data-defined person records on all nonblank, primary paper returns. A data-defined person record had at least two entries that met

specified completion criteria for the following items⁷: name, relationship, sex, age/date of birth, Hispanic origin, race, and ancestry.

2.8 Quality Assurance Procedures

Quality assurance procedures were applied to the analysis and preparation of this report. The procedures encompassed data processing, data verification, factual content, technical writing, relevance, technical review and clearance, as appropriate. A description of the procedures used is provided in the “Handbook for the Quality Process for 2010 Census Test Evaluations.”

3. LIMITATIONS

3.1 Population of Inference

The purpose of this experiment was to determine response to the bilingual form for a national sample of housing units in mailout/mailback areas. This experiment was not designed to study the efficacy of the bilingual form in targeted areas with a large portion of the Spanish-speaking population. However, the sample was equally allocated to two strata, one of which contained blocks with a high concentration of the non-White or Hispanic populations. Therefore, the estimates that we compute by strata may indicate the efficacy of the form in these areas but still give no indication of impact on Spanish-speaking populations. Note that there were plans for a targeted bilingual form experiment in the 2006 site test, but due to budgetary issues this research was canceled.

3.2 Confounding Factors

There were certain inherent design issues for the bilingual form that triggered variation from the English (control) design. The following list denotes all known variations from the English control form to the bilingual form:

- The advance letter for the bilingual form panel had the dual language swim lane design.
- The cover letter, typically a separate document in the initial and replacement questionnaire mailings, had the swim lane design and formed the first page of the questionnaire booklet for the bilingual form.
- The bilingual form was not folded prior to mailing, so it was sent in a larger envelope than the English form, which was folded.
- The bilingual form required a reminder letter in lieu of a postcard (due to space constraints), and this letter had the swim lane design.
- The Internet invitation was included in the cover letter, which formed the first page of the questionnaire booklet for the bilingual panel. For the English panel, the invitation was included in a cover letter, which was separate from the questionnaire. An additional Internet invitation was included above the roster on the first page of the English form; the bilingual form had no such additional invitation. This difference could have impacted the selection of response mode between the English and bilingual panel respondents.

⁷ For more information about the criteria used to determine data-defined status, please see the memorandum from Stapleton (2004).

- Due to spacing issues, the relationship question response categories were different between the bilingual form and the control panel (see appendices A and B for differences). Therefore, comparisons were not made across panels for the relationship item.
- Since various treatments were combined in panels to limit costs, the panel that was used as the control for this evaluation had a different age/date of birth order than the bilingual form. Therefore, comparisons were not made across panels for these items.

We are unable to determine what role, if any, these factors played in any significant differences between the bilingual form and the English form. That is, we evaluated the bilingual form with all of these changes as a package, and are unable to estimate the effects of any one factor.

3.3 Census Test Environment

Note that results from a census test may differ from results in an actual decennial census due to differences in media attention, advertising and scope. We cannot determine whether public reaction to the bilingual form would be different in a true decennial environment.

3.4 Natural Disasters

The 2005 NCT initial questionnaire mailing coincided with Hurricane Katrina. Mail service to some areas around the gulf coast in Louisiana, Mississippi, and Alabama was disrupted or suspended during the data collection period. These areas contained 3,200 of our 420,000 sample cases across all panels in the 2005 NCT. We treated these cases as UAA, and as such, they were excluded from the analyses⁸.

4. RESULTS

4.1 Self-Response Rate Comparisons

This section of the report examines the effect of the bilingual form on response to the census test, as compared to the standard English form. Table 1 below contains the self-response rates for the bilingual form panel and the English control panel at a national level and by strata. Additionally, the rates are further studied by response mode. Breaking out the self-response rates in this manner allows us to examine whether receipt of the bilingual form influenced the decision to respond by paper or Internet, keeping in mind the differences in how the response modes were offered (see sections 2.2 and 3.2).

⁸ For more detailed information on the cases affected by Hurricane Katrina, see Zajac (2005) and Tancreto (2006).

Table 1. Self-Response Rates and Standard Errors (in percent) by Language and Response Mode at the National Level and by Strata

Treatment	National			High Non-White or Hispanic Concentration			Low Non-White or Hispanic Concentration		
	Total	Paper	Internet	Total	Paper	Internet	Total	Paper	Internet
Bilingual Form	62.0	55.5	6.5	47.0	42.7	4.3	68.8	61.3	7.5
English Form	60.8	53.3	7.5	44.2	39.5	4.7	68.4	59.6	8.8
Difference	1.1* (0.61)	2.2*** (0.60)	-1.0*** (0.33)	2.8*** (0.87)	3.2*** (0.81)	-0.4 (0.35)	0.4 (0.79)	1.7** (0.79)	-1.3*** (0.46)

* Indicates difference is statistically significant at the $\alpha=.10$ level.

**Indicates difference is statistically significant at the $\alpha=.05$ level.

*** Indicates difference is statistically significant at the $\alpha=.01$ level.

As previously mentioned, most results in this report focus on paper response, since there was no Spanish version of the Internet to complement the bilingual form. Looking at the paper self-response rates in Table 1, we see that the bilingual panel achieved significantly higher response (by 2.2 percentage points) compared to the English panel.

Not surprisingly, this increase in paper response was even more evident in the high non-White or Hispanic concentration stratum, suggesting that the bilingual form was particularly effective in areas with high proportions of the Hispanic or non-White populations. Somewhat surprisingly, we found that the bilingual form also showed a significant increase in paper response in areas with a low concentration of the non-White or Hispanic populations relative to the English form. We would have expected no effect in these areas or potentially even a negative effect if there were opposition to the use of a bilingual form. Therefore, this finding may suggest that there is no negative effect of offering a bilingual form in areas that have a heavy concentration of non-Hispanic whites⁹.

Please note that there are limitations associated with these paper response rate comparisons, as well as the Internet response rates. These limitations stem from differences in the administration of the Internet across the two panels, which may have influenced respondents in choosing a particular response mode. The first difference in Internet administration pertains to the invitation to use the Internet. While both panels (bilingual and English only) offered the option of responding by Internet, the invitation to use the Internet was located in different places between the English and bilingual panels. Both panels provided the Internet option in the cover letter¹⁰, but the English panel also provided the option again on the first page of the questionnaire (above the first question), while the bilingual form did not¹¹. Secondly, this test did not provide a

⁹ This finding is based on the distribution of bilingual forms to 10,000 housing units across the United States. It is possible that public attention would impact this finding if bilingual forms were distributed on a larger scale.

¹⁰ Recall that the cover letter for the bilingual panel formed the first page of the bilingual questionnaire, whereas the cover letter for the English panel was contained in the mailing package as a separate document.

¹¹ Although the bilingual form did not have a second Internet invitation on the first page of the questionnaire, this difference may not have had a substantial impact on self-response for that panel. Item nonresponse analysis shows

Spanish version of the Internet, so the bilingual cover letter noted the disclaimer “English only” next to the invitation to respond by Internet.

We cannot conclusively determine the impact of these differences on paper response. We would expect that having one less Internet invitation and no Spanish Internet would lower Internet response¹²; however, it is unclear as to the impact on paper response. We could imagine lower paper response if those who would have used the Internet chose not to respond at all. Conversely, we could see an increase in paper response if those who would have used Internet substituted the paper form in its place.

Therefore, to the extent that these differences in Internet administration influenced response mode selection, the results from the paper response rate comparisons, as well as Internet rate comparisons, may not be valid. That is, the paper response rates (and related results) may be overstated, while the Internet rates may be underestimated, especially under the substitution theory. Assuming people who would have responded by Internet substituted paper, we can study total response rates (paper and Internet) to gauge the effect of the bilingual form. Similar to the paper response rate, the difference in total self-response rates was also significantly higher for the bilingual panel compared to the English only panel, both nationally (by 1.1 percentage points) and for the high non-White or Hispanic concentration stratum (by 2.8 percentage points). Paper response for the low non-White or Hispanic stratum was significantly higher for bilingual than English, but this difference did not remain when combining both paper and Internet returns. Again, the total response rate estimates may not be meaningful due to survey administration differences that could have affected both paper and Internet response rates.

Finally, we looked at self-response by form (initial or replacement questionnaire) for each panel (table not shown). Of those who responded to the bilingual paper form, 83.2 percent used the initial questionnaire and 16.8 percent responded using the replacement questionnaire. For the English panel, 83.7 percent responded using the initial questionnaire and 16.3 percent responded using the replacement questionnaire. There was no difference in self-response between the two panels for the initial and replacement questionnaires.

4.2 Bilingual Form Response Patterns

Next, we studied response patterns for those who responded via the bilingual form. We set out to get a better understanding of how people were using the form. First, we studied the percentage of respondents that used each column. We found that 2.7 percent of bilingual returns predominantly used the Spanish language column to complete the form, while the remaining 97 percent used the English column. In areas with a high concentration of the non-White or Hispanic populations, 8.3 percent of bilingual returns used the Spanish column whereas less than 1 percent of bilingual returns from areas with low concentrations of the non-White or Hispanic populations used the Spanish column.

higher household item nonresponse for the bilingual form, which may indicate that those respondents missed the first page of the questionnaire.

¹²These differences did not result in a significantly lower Internet response rate for the bilingual panel in the high non-White and Hispanic concentration stratum, where we might expect the absence of a Spanish Internet form to have the most effect.

We also considered whether households used both language columns in responding (i.e., lane jumping). For those that used both columns, we studied how they incorporated both language columns in responding (e.g., reported the same person twice, once in each language column, etc.). There were 45 forms (less than 1 percent of bilingual paper returns) where respondents provided data in both language columns. Of the 45 forms, just under half of the respondents (n = 19) appeared initially confused, as they completed only the first few items (i.e., household count, name, sex) in both columns but then provided all remaining answers in one column¹³. Also, a little less than half of the respondents (n = 19) seemed to randomly switch language columns for various questions. That is, there was no particular pattern of column use for these forms. The balance (n = 7) used different language columns for different persons within a household. For these cases, there may have been more than one respondent, where each respondent exhibited a personal preference in choosing a language column.

The fact that less than 1 percent of bilingual form respondents used more than one column in responding seems to suggest that respondents are using the form in the way it was intended to be used. That is, most households chose to respond in one language. The seven cases where different languages were used for each person is also one of the anticipated uses of the form. This form allows household members to provide information for themselves in the language that is most comfortable for them.

We also looked at lane jumpers by Hispanic origin of Person 1 and found that 84.3 percent of forms with lane jumping had Person 1 Hispanic. We also looked at lane jumpers by language column use and found that 14.2 percent of households that predominantly used the Spanish language column were lane jumpers.

4.3 Item Nonresponse Rate Analysis

4.3.1 Item Nonresponse Rates by Panel

Next we considered whether the language of the paper form influenced the presence of a response to the questions. Item nonresponse is important to study, as it is one indicator of the extent to which a particular item may be subject to nonresponse bias. For this analysis, we calculated the item nonresponse rates for five person level data items (sex, Hispanic origin, race, ancestry, and overcount) and four housing unit level data items (household count, tenure¹⁴, undercount and telephone number). Note that we cannot compare rates for age/year of birth and relationship between the bilingual and English forms due to differences in item wording (see appendices A and B). When computing rates for the bilingual form, total item nonresponse took into account both language columns; that is, we deemed an item missing if there was no response in the English column and there was no response in the Spanish column. Table 2 contains national item nonresponse rates for the bilingual form and English form.

¹³ Note that we did not see a trend of reporting the same person twice in the English and Spanish columns. However, completing the first few person level items could lead to “extra” data-defined persons being included in processing, which could impact coverage.

¹⁴ Note that there is a slight wording difference for the tenure response options across forms, however, we do not expect this difference to impact item nonresponse rates.

Table 2. Item Nonresponse Rates and Standard Errors (in percent) by Language Treatment

Item	English Form	Bilingual Form	Difference (Bilingual – English)	
	Estimate	Estimate	Estimate	Std. Err.
Household Items				
Household count	1.1	2.7	1.6***	0.27
Tenure	1.5	3.5	2.0***	0.29
Undercount	7.3	10.4	3.1***	0.51
Telephone number	8.1	10.0	1.9***	0.53
Person Items				
Sex	0.6	0.6	0.0	0.09
Hispanic origin	2.6	4.5	1.9***	0.26
Race	3.6	3.9	0.3	0.31
Ancestry	13.2	12.5	-0.7	0.65
Overcount	1.1	1.0	0.0	0.14

* Indicates difference is statistically significant at the $\alpha=.10$ level.

**Indicates difference is statistically significant at the $\alpha=.05$ level.

*** Indicates difference is statistically significant at the $\alpha=.01$ level.

Table 2 shows item nonresponse rates for the bilingual panel were significantly higher than the English only panel for household count, tenure, undercount, telephone number and Hispanic origin¹⁵. Note that while undercount and telephone number had higher item nonresponse rates for the bilingual panel, these items are generally used for operational purposes and are not publicly reported. Item nonresponse rates for sex, race, ancestry and overcount were not statistically different across treatments.

The difference in item nonresponse rates for all household items was somewhat surprising so we looked further at the percent of households with all household data missing. The bilingual panel had a significantly higher percent of forms with all household level items missing (2.1 percent) compared to the English panel (0.4 percent). This may suggest that the increase in item nonresponse for the household items on the bilingual form is more of a bilingual form design issue rather than an item nonresponse issue (i.e., the bilingual form design allowed respondents to miss the first page).

The first page of the bilingual form that contains the household items has less “white space” than the English only form since the residency rules are included twice (once in English and once in Spanish) on the page (see appendices A and B). Respondents may have quickly looked at this page and dismissed it as simply more instructions. Also, preliminary results from cognitive testing of the bilingual form showed that some Spanish-speaking respondents tended to look to the left side of the form (the English column) on each page before locating the Spanish column

¹⁵ We studied whether we could impute Hispanic origin from the ancestry question when the Hispanic origin question was blank. We found that the item nonresponse rate for Hispanic origin dropped by 70% for both panels using the imputed ancestry data; however, the bilingual panel still exhibited significantly higher item nonresponse than the English panel.

(Caspar *et al.*, 2006). Thus, it is possible that some respondents may not have noticed the Spanish column until getting further into the questionnaire.

In addition to the potential forms design issue, there are several other factors that may have contributed to the higher prevalence of missing household data on the bilingual form. For instance, this increase in household item nonresponse may be a product of the population responding to the bilingual form. As we show in section 4.4 of this report, the bilingual form increases the percent of Hispanics included on the returns compared to the English only form. Moreover, the increase in household item nonresponse for the bilingual form may also be related to question translation since there was limited pre-testing conducted on this form. Finally, as suggested above, early cognitive research results suggest that the Spanish column may be initially overlooked (Caspar *et al.*, 2006). Thus, there are quite a few factors that may be driving the item nonresponse differences between the bilingual and English forms.

Finally, we also looked at the percent of *sampled housing units* (excluding UAA cases) with all household level items missing for the bilingual panel¹⁶. We wanted to get a better understanding of the tradeoff between the increase in missing household data and the increase in paper self-response rates. Thus, we compared the percent of sampled housing units with all household level data missing (1.1 percent) against the percentage point increase in paper response to the panel (2.2 percentage points). When examining these results, we see that there is still a 1.1 percentage point net gain in overall paper response when taking into account the rate of missing household data.

We also compared item nonresponse rates across the treatments for each stratum separately in Table 3.

¹⁶ The full sample item nonresponse rates for the bilingual panel household items are: household count, 1.5 percent; tenure, 1.9 percent; undercount, 5.8 percent and telephone, 5.7 percent.

Table 3. Item Nonresponse Rates and Standard Errors (in percent) by Language Treatment and Strata

Item	High Non-White or Hispanic Concentration				Low Non-White or Hispanic Concentration			
	English Form	Bilingual Form	Difference (Bil – Eng)		English Form	Bilingual Form	Difference (Bil – Eng)	
	Est.	Est.	Est.	Std. Err.	Est.	Est.	Est.	Std. Err.
Household Items								
Household count	1.6	4.0	2.5***	0.46	0.9	2.3	1.4***	0.32
Tenure	2.3	4.9	2.6***	0.50	1.2	3.1	1.8***	0.35
Undercount	9.2	13.2	4.0***	0.88	6.8	9.5	2.7***	0.62
Telephone number	9.3	11.0	1.7**	0.77	7.7	9.7	1.9***	0.66
Person Items								
Sex	0.8	0.9	0.2	0.21	0.5	0.5	-0.1	0.10
Hispanic origin	3.8	5.4	1.6***	0.43	2.2	4.2	2.1***	0.31
Race	8.3	10.2	1.9**	0.90	2.1	1.7	-0.4	0.27
Ancestry	12.0	11.2	-0.9	0.91	13.6	13.0	-0.6	0.82
Overcount	1.5	1.8	0.3	0.34	0.9	0.7	-0.2	0.15

* Indicates difference is statistically significant at the $\alpha=.10$ level.

**Indicates difference is statistically significant at the $\alpha=.05$ level.

*** Indicates difference is statistically significant at the $\alpha=.01$ level.

As for panel comparisons by strata, the results were similar to the national results with all household items and Hispanic origin being significantly higher for the bilingual panel compared to the English only panel, for both strata. Additionally, within the high stratum, the race item nonresponse rate was significantly higher for the bilingual panel as compared to the English only panel. This difference in item nonresponse for the race item was not surprising since there were more Hispanics reported on the bilingual form (see section 4.4) and Hispanics tend to omit race more than non-Hispanics (Martin *et al.*, 2004; del Pinal, 2003).

Results from the 1993 Spanish Forms Availability Test suggested that Hispanics were more likely to omit information on the Spanish form (bilingual or separate form) compared to Hispanics who responded to the English form (Corteville, 1994; de la Puente *et al.*, 1994a). Therefore, we compared item nonresponse rates across the forms for Hispanic and non-Hispanic respondents to test whether this result was replicated in this experiment (see Table 4). We used the characteristics of Person 1 to describe the respondent. Admittedly, past research suggests that Person 1 is not the respondent roughly 30 percent of the time (DeMaio *et al.*, 1990)¹⁷.

¹⁷ We also ran the item nonresponse rates for the person items by Hispanic origin of each person (not just Person 1). We also found no significant differences in item nonresponse between panels for the person items for Hispanics and non-Hispanics using this approach.

Table 4. Item Nonresponse Rates¹⁸ and Standard Errors (in percents) by Language Treatment and Person 1 Hispanic/Non-Hispanic¹⁹

Item	Person 1 Hispanic				Person 1 Non-Hispanic			
	English Form	Bilingual Form	Difference (Bil – Eng)		English Form	Bilingual Form	Difference (Bil – Eng)	
	Est.	Est.	Est.	Std. Err.	Est.	Est.	Est.	Std. Err.
Household Items								
Household count	1.0	2.4	1.5**	0.74	1.0	1.6	0.6***	0.22
Tenure	2.4	4.1	1.6*	0.99	1.2	2.2	1.0***	0.25
Undercount	12.4	15.8	3.3*	1.96	6.4	8.6	2.2***	0.51
Telephone number	8.1	9.5	1.4	1.51	7.8	8.6	0.8	0.53
Person Items								
Sex	0.8	0.8	0.1	0.40	0.5	0.5	0.0	0.09
Race ²⁰	27.1	26.1	-1.0	2.49	1.0	1.1	0.1	0.13
Ancestry	14.8	13.4	-1.4	2.02	12.6	11.8	-0.8	0.69
Overcount	2.9	2.2	-0.7	0.76	0.7	0.7	0.0	0.10

* Indicates difference is statistically significant at the $\alpha=.10$ level.

**Indicates difference is statistically significant at the $\alpha=.05$ level.

*** Indicates difference is statistically significant at the $\alpha=.01$ level.

Table 4 shows that, for Hispanic and non-Hispanic respondents alike, item nonresponse for household count, tenure and undercount are significantly higher for the bilingual panel as compared to the English only panel. Again, this could be a function of the form design in completing the household data on the bilingual form, as previously mentioned. No other differences in item nonresponse were found by Hispanic origin of Person 1. These results provide limited support for the SFAT evaluation finding that Hispanics are more likely to omit data on a bilingual form than Hispanics who use an English form (de la Puente *et al.*, 1994a).

For both panels, the item nonresponse rate for the race question is considerably higher for Hispanic respondents (27.1 percent for English form and 26.1 percent for bilingual form) than for Non-Hispanic respondents (1.0 percent for the English form and 1.1 percent for the bilingual form). This trend is consistent with past findings that Hispanics are less inclined to answer the race question than non-Hispanics (Martin *et al.*, 2004; del Pinal, 2003). That is, some Hispanics find reporting on both Hispanic origin and race to be redundant or they simply do not apply the race categories to themselves (Martin *et al.*, 1990).

¹⁸ Note that we found a slight correlation between missing data for Hispanic origin and missing data for other items, suggesting that people who leave Hispanic origin blank are more likely to leave other items blank as well. Therefore, item nonresponse rates in Table 4 will be slightly lower than expected, since respondents missing Hispanic origin are excluded in the table.

¹⁹ Sample sizes: Person 1 Hispanic/English form: $n = 1,371$; Person 1 Hispanic/Bilingual form: $n = 512$; Person 1 Non-Hispanic/English form: $n = 11,937$; Person 1 Non-Hispanic/Bilingual form: $n = 4,046$.

²⁰ The lack of difference in race item nonresponse rates in Table 4 may also be evidence that the estimates are low. That is, the race item nonresponse rate was significantly higher for the bilingual panel overall and for areas with a high concentration of non-White or Hispanic populations, however, it was not significantly different across panels by Hispanic origin for Person 1.

4.3.2 Item Nonresponse Rates – Bilingual Form Only

We then examined item nonresponse rates between the two language columns on the bilingual form (see Table 5 below). For this analysis, we calculated the item nonresponse rates for seven person level data items (age/year of birth, sex, relationship, Hispanic origin, race, ancestry, and overcount) and four housing unit level data items (household count, tenure, undercount and telephone number). We assigned persons or households to each language column to compute rates based on the language column that was used most heavily²¹. For example, if Person 2 provided all data in Spanish with the exception of sex, which was provided in the English column, he or she was counted in the Spanish column and was included as item nonresponse for the sex question.

Table 5. Item Nonresponse Rates and Standard Errors (in percent) by Language Column on the Bilingual Form²²

Item	English Column	Spanish Column	Difference (Spanish – English)	
	Est.	Est.	Est.	Std. Err.
Household Items				
Household count	2.6	7.8	5.3 ^{***}	1.97
Tenure	3.3	9.3	6.0 ^{***}	2.09
Undercount	10.0	24.8	14.8 ^{***}	3.46
Telephone number	10.0	9.9	-0.1	2.27
Person Items				
Age/Year of birth	0.8	1.4	0.6	0.59
Sex	0.6	1.3	0.7	0.96
Relationship (per 2+)	0.8	2.1	1.3	0.90
Hispanic origin	4.5	4.7	0.2	1.13
Race	2.9	31.4	28.5 ^{***}	3.94
Ancestry	12.6	12.0	-0.6	2.84
Overcount	0.9	5.3	4.5 ^{***}	1.59

* Indicates difference is statistically significant at the $\alpha=.10$ level.

**Indicates difference is statistically significant at the $\alpha=.05$ level.

*** Indicates difference is statistically significant at the $\alpha=.01$ level.

Table 5 shows item nonresponse rates for household count, tenure, undercount, race and overcount were significantly higher for the Spanish column as compared to the English column. The undercount item nonresponse rate for the Spanish column was considerably higher (by 14.8 percentage points) than the English column. Again the household item results were somewhat unexpected, so we looked at the percent of bilingual forms with all household items missing by language. The respondents that used the Spanish column had a significantly higher proportion of respondents that had all household items missing (5.9 percent) than the respondents who used the English column (2.0 percent). As noted above, there are many potential reasons for this difference, including question translation and form design.

²¹ If both columns were completed equally, we randomly assigned the person or household to one of the columns.

²² Sample sizes: English column: n = 4,599; Spanish column: n = 178.

Looking at the person items, we see that the race item nonresponse rate was significantly higher for the Spanish column (by 28.5 percentage points) than the English column. Note that the vast majority of people who used the Spanish column were Hispanic (see section 4.4), and Hispanics are less likely to answer race, as previously mentioned. Additionally, cognitive testing of the bilingual form with Spanish-speaking respondents revealed that the note under the race question (“for this census, Hispanic origins are not races”) might have inadvertently served as a skip instruction for some Spanish-speaking respondents (Caspar *et al.*, 2006).

We also noticed an increase in item nonresponse for the coverage overcount question. The intent of the coverage overcount question was to detect whether each person listed on the form should be counted at another place. The noticeable increase in item nonresponse for the overcount question may indicate that this question is sensitive to Spanish-speaking respondents.

Finally, we studied item nonresponse between the two language columns on the bilingual form for Hispanics. Table 6 presents the item nonresponse rates by language column for cases where Person 1 (i.e., respondent) was Hispanic²³. Item nonresponse rates are not shown for Person 1 non-Hispanic due to the fact that too few non-Hispanics filled out the Spanish column, so we could not make comparisons between the columns.

Table 6. Item Nonresponse Rates and Standard Errors (in percent) by Language Column for Person 1 Hispanic on the Bilingual Form²⁴

Item	Person 1 Hispanic			
	English Column	Spanish Column	Difference (Spanish – English)	
	Est.	Est.	Est.	Std. Err.
Household Items				
Household count	1.2	6.0	4.8***	1.79
Tenure	3.3	6.0	2.7	2.03
Undercount	13.1	22.1	9.0**	3.89
Telephone number	10.7	6.6	-4.0	2.67
Person Items				
Age/Year of birth	1.4	1.5	0.1	0.82
Sex	0.6	1.4	0.8	1.04
Relationship (per 2+)	2.1	1.5	-0.6	1.20
Race	22.8	32.6	9.7**	4.70
Ancestry	13.9	12.3	-1.6	3.55
Overcount	0.9	4.9	4.0**	1.59

* Indicates difference is statistically significant at the $\alpha=.10$ level.

**Indicates difference is statistically significant at the $\alpha=.05$ level.

*** Indicates difference is statistically significant at the $\alpha=.01$ level.

Table 6 shows that item nonresponse rates for household count, undercount, race and overcount were significantly higher for the Spanish column than for the English column for Hispanic

²³ We also ran person item nonresponse rates by Hispanic origin of each person (not just Person 1). We found the same significant differences (race and overcount), while all other person items were not significant.

²⁴ Sample sizes: English column: n = 352; Spanish column: n = 160.

respondents. These results provide some support for the SFAT result that showed that item nonresponse rates were higher for Spanish language than English language for Hispanic respondents, specifically for household items and race (de la Puente *et al.*, 1994a). Note that caution should be used in interpreting these results due to small cell sizes in both columns (see footnote for sample counts).

4.4 Comparisons of Demographic Characteristics

The 2005 NCT sample was selected such that each panel contained a random sample of housing units from the mailout/mailback universe. Therefore, we would expect that the persons for whom data are provided would be demographically similar across the panels, unless the offer of a bilingual form influenced the self-selection of those that chose to respond.

We hypothesized that the bilingual form would be more appealing than the English form to certain people, especially those who are primarily Spanish-speaking. This analysis helped determine whether the people listed on the bilingual form differed from the people listed on the English form. Again, the data for this comparison came solely from the paper returns for both panels.

Table 7. Demographic Characteristics and Standard Errors (in percents) by Language Treatment

Item	English Form	Bilingual Form	Difference (Bilingual – English)	
	Est.	Est.	Est.	Std. Err.
Average household size	2.4	2.5	0.1**	0.04
Male	47.8	46.5	-1.3***	0.41
Hispanic	10.5	12.0	1.5***	0.54
Race				
White	80.1	79.5	-0.6	0.67
Black	9.0	9.1	0.0	0.44
American Indian/Alaska Native	0.5	0.6	0.1	0.13
Asian	4.2	4.5	0.2	0.37
Native Hawaiian/Pacific Islander	0.3	0.3	0.0	0.10
Some Other Race	3.9	4.5	0.6	0.39
Two or more races	2.0	1.7	-0.3	0.21

* Indicates difference is statistically significant at the $\alpha=.10$ level.

**Indicates difference is statistically significant at the $\alpha=.05$ level.

*** Indicates difference is statistically significant at the $\alpha=.01$ level.

Table 7 shows the distribution of demographic characteristics for each panel. In general, there are few demographic differences between English form returns and bilingual form returns, with three exceptions. Household size and the proportion of Hispanic persons were both significantly higher for the bilingual form. The English form had more males compared to the bilingual form.

The increase in the percent of Hispanics on the bilingual form is noteworthy, since Hispanics have been notably undercounted in previous censuses (U.S. Census Bureau, 2004). This finding

also justifies the observed increase in household size. With more Hispanics responding to the bilingual form, the larger household size is consistent with findings from Ramirez *et al.* (2002), which shows that Hispanics live in larger households than non-Hispanics. However, the finding that there are more males reported on the English form than the Bilingual form contradicts previous research by Ramirez (2004), which suggests that males outnumber females in the Hispanic community. We have no immediate explanation for this finding.

Table 8 below shows the distribution of demographic characteristics for each language column on the bilingual questionnaire²⁵. Note that race categories for Black, American Indian/Alaska Native, Asian, Native Hawaiian/Pacific Islander and Two or more races were removed from the table due to small cell counts and disclosure avoidance issues.

Table 8. Demographic Characteristics and Standard Errors (in percents) by Language Column on the Bilingual Form

Item	English	Spanish	Difference	
	Column Est.	Column Est.	(Spanish – English) Est.	Std. Err.
Average household size	2.5	3.8	1.3**	0.17
Renters	22.9	39.3	16.3***	4.39
Average age	41.5	31.1	-10.5***	.96
Male	46.3	51.1	4.8**	1.45
Hispanic	8.7	97.1	88.4***	.96
Race				
White	80.1	57.2	-22.8**	5.53
Some Other Race	3.4	41.6	38.1***	5.42

* Indicates difference is statistically significant at the $\alpha=.10$ level.

**Indicates difference is statistically significant at the $\alpha=.05$ level.

*** Indicates difference is statistically significant at the $\alpha=.01$ level.

The data suggest several differences between English column responses and Spanish column responses. A significantly higher proportion of people in the Spanish column were renters, male, Hispanic and of Some Other Race. These persons were also from larger households and younger than those in the English column.

Table 8 shows that 97.1 percent of the people on the bilingual form that used the Spanish column were Hispanic. This result was not surprising, nor were the other demographic results shown in Table 8, as they have been shown to be correlated with being Hispanic. That is, past research (Ramirez, 2004) supports the notion that Hispanics tend to be younger and renters than non-Hispanics. This research also shows that males outnumber females in the Hispanic population (Ramirez, 2004). Furthermore, household size results are consistent with findings from Ramirez *et al.* (2002), which show that Hispanics live in larger households than non-Hispanics. And finally, the race results are also consistent with past research, in that, Hispanics are more likely to report their race as either “White” or “Some Other Race” (when the question is asked with this particular wording) compared to non-Hispanics (Grieco *et al.*, 2000).

²⁵ For the person items, language column was assigned to each person based on the column that was used most heavily in reporting data for that person.

4.5 Public Reaction to the Bilingual Form

There was no formal evaluation to scientifically measure the reaction of respondents to the bilingual form in the 2005 NCT. Therefore, we cannot infer from this experiment what public reaction would be if the bilingual form were included in a decennial census. As an attempt to get some indication, the Decennial Management Division (DMD) coordinated an effort to document anecdotal information on the public's reaction to receiving a bilingual census form. In early September 2005, the External Liaison Branch of DMD widely distributed an email message to regional offices and other Census Bureau divisions providing guidance on how to answer respondent inquiries concerning the 2005 NCT. Most importantly, this message included a request for these areas to document any inquiries related to the 2005 NCT Bilingual Questionnaire. Additionally, DMD combed through bilingual form returns that included written correspondence (i.e., comments written on the questionnaire) to provide insight on reactions to the use of dual languages.

This investigation resulted in less than ten instances of communication regarding the bilingual questionnaire. In these inquiries, almost all respondents wondered why they received an English and Spanish questionnaire. DMD encountered no negative public feedback concerning the 2005 NCT bilingual questionnaire.

While we did not scientifically measure public reaction to the form, we can say that the data for this 10,000 housing unit sample showed no sign of backlash since we did not see any drop in self-response rates, even in areas that were heavily concentrated with non-Hispanic whites. However, it is important to note that this experiment was limited to 10,000 housing units across the country, and we have no way of knowing what the reaction would be if the form were distributed on a much larger scale.

5. CONCLUSIONS AND FUTURE CONSIDERATIONS

Conclusions

The data from the 2005 National Census Test show that a bilingual form significantly increases self-response nationally, and more specifically, in areas where there is a high concentration of non-White or Hispanic populations. Item nonresponse rates for the bilingual form were higher for most household level items, which may suggest a possible forms design issue or may be a product of the population responding (i.e., more Hispanic respondents). In addition, there were a minimal amount of bilingual form lane jumpers, which may suggest there was little confusion with the dual language design of the bilingual form. Finally, the bilingual form had larger households, more Hispanic persons and fewer males compared to the English form.

Future considerations

An important issue that needs to be considered in offering a bilingual form is that of data processing. As a result of dual language design, the bilingual data may require different processing standards and procedures if implemented in a decennial census.

For the 2005 NCT, the bilingual data were captured such that there were two records for each household and for each person (one record for English and one for Spanish). This is not the format used for standard capture data, where there is generally one record per household and one per person. Also, evaluation staff rather than the processing area conducted data processing for the 2005 NCT. As a result, there was no system test to see how the bilingual data could be integrated into the standard data processing system in terms of format as well as edits and imputations.

Additionally, some respondents provided data in both languages (i.e., lane jumpers). These data were clerically reviewed to determine which language record was the primary record. Also, in some cases, valid data needed to be transferred to the primary record. (For example, a respondent completed the form in English, except for race, which was completed in Spanish. The race data needed to be included in the English record so as not to be deemed missing.)

Another important issue that needs to be considered in offering a bilingual form in the future is that of public reaction in a decennial census environment. We have no indication of any backlash (both from data and feedback) from this test for a national sample of 10,000 housing units; however, we cannot predict public reaction if the form is distributed to a larger audience in the 2010 census.

ACKNOWLEDGEMENTS

We would like to acknowledge the following people for their assistance with this evaluation.

Mike Bentley for QA support, Joan M. Hill and Henry Woltman for their review and technical advice, Kathleen Styles, Jane Ingold, and Carol Pendleton for collecting and summarizing public reaction and Patricia Goerman, for sharing early results from cognitive testing.

REFERENCES

- Bentley, Michael (2005) "Sample Specifications for the 2005 National Census Test", U.S. Census Bureau Internal Memorandum from Killion to Longini, version 5.0, March 9.
- Caspar, Rachel, Goerman, Patricia, McAvnichey, Georgina, Quiroz, Rosanna, and Sha, Mandy (2006) "Census Bilingual Questionnaire Research -- Final Round 1 Report", Report submitted by RTI International to U.S. Census Bureau, July 7.
- Corteville, Jeffrey (1994) "Spanish Forms Availability Test Completion Rate Evaluation", U.S. Census Bureau Internal Memorandum from Thompson to Miskura, DSSD Memorandum Series #D-8, July 18.
- De la Puente, Manuel and Wobus, Peter (1994a) "Final Report of Results From Item Nonresponse Analysis for the Spanish Language Forms Availability Test", U.S. Census Bureau, Center for Survey Methods Research, September.
- De la Puente, Manuel and Wobus, Peter (1994b) "An Item Nonresponse Analysis and Log-Linear Analysis of the Spanish Language Forms Availability Test", *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, August, pp. 583-588.
- De la Puente, Manuel and Wobus, Peter (1995) "Results from Telephone Debriefing Interviews: The Census Bureau's Spanish Language Forms Availability Test", paper presented at the American Association for Public Opinion Research, May.
- Del Pinal, Jorge (2003) "Race and Ethnicity in Census 2000", Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report Series, Number 9.
- DeMaio, Theresa J. and Bates, Nancy A. (1990) "Who Fills Out the Census Form?", *Proceedings of the Survey Research Methods Section of the American Statistical Association*, pp. 584-589.
- Grieco, Elizabeth M., and Cassidy, Rachel C. (2001) "Overview of Race and Hispanic Origin", Census 2000 Brief, U.S. Census Bureau.
- Martin, Elizabeth, Sheppard, Dave, Bentley, Michael, and Bennett, Claudette (2004) "Results of the 2003 National Census Test of Race and Hispanic Questions", U.S. Census Bureau Internal Report, September 7.
- Ramirez, Roberto R., and de la Cruz, G. Patricia (2002) "The Hispanic Population in the United States: March 2002", Current Population Reports, U.S. Census Bureau.
- Ramirez, Roberto R. (2004) "We the People: Hispanics in the United States", Census 2000 Special Reports, U.S. Census Bureau.
- Rothhaas, Cynthia (2004a) "Definition of Undeliverable As Addressed Flag for the 2005 National Census Test", U.S. Census Bureau Internal Memorandum for the Record, 2005 Analysis Team, March 15.

Rothhaas, Cynthia (2004b) “Definition of Blank Forms for the 2005 National Census Test -- revised”, U.S. Census Bureau Internal Memorandum for the Record, 2005 Analysis Team, May 24.

Rothhaas, Cynthia (2004c) “Duplicate Return Rule for the 2005 National Census Test -- revised”, U.S. Census Bureau Internal Memorandum for the Record, 2005 Analysis Team, November 3.

Stapleton, Courtney (2004) “Definition of Data-Defined Persons for the 2005 National Census Test – Revised Draft”, U.S. Census Bureau Internal Memorandum for the Record, 2005 Analysis Team, May 31.

Tancreto, Jennifer G., and Bouffard, Julie A. (2005) “Final Draft of the Study Plan for 2006 Language Program Mail Strategies Evaluation”, 2010 Census Test Memoranda Series, 2005 National Census Test Chapter, Number 17, August 15.

Tancreto, Jennifer G. (2006) “2005 National Census Test: Global Test Measure and Issues during Production”, U.S. Census Bureau Internal Memorandum from Tancreto to Boone and Ingold, DSSD 2005 Memorandum Series #E-2, January 25.

U.S. Census Bureau (2004), “Coverage Measurement From the Perspective of March 2001 Accuracy and Coverage Evaluation”, February.

U.S. Census Bureau (2005), “2005 National Census Test Panel Design Proposal”, U.S. Census Bureau Internal Memorandum Baseline 4.0, May 24.


Zajac, Kevin (2005) Modifications to 2005 National Census Test Analysis Files for Hurricane Katrina Areas – *revised: 12/21/05*”, U.S. Census Bureau Internal Memorandum from 2005 Analysis Team to Imel, December 21.

Appendix A

2005 National Census Test Bilingual Form (Panel 20)

2005 National Census Test

Censo Nacional de Prueba de 2005

**UNITED STATES DEPARTMENT OF COMMERCE**
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20235-0001
OFFICE OF THE DIRECTOR

August 29, 2005

Dear Resident:

The United States Constitution requires a census of the United States every 10 years. To prepare for the 2010 Census, the U.S. Census Bureau is conducting the 2005 National Census Test. The purpose of this test is to develop new methods that will make the 2010 Census easier, more convenient, and less costly for taxpayers.

There are two ways to provide census information for all people living at this address. Please choose **ONLY** one:

OPTION 1 > Go to www.census.gov/census2005 to complete your questionnaire on-line. (English only)

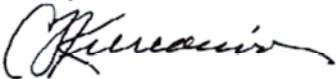
OR

OPTION 2 > Complete and mail back the enclosed questionnaire in the postage-paid envelope provided.

Your assistance in helping us improve the census is important. Results from the 2010 Census will be used to help each community get its fair share of federal funding.

Your answers are required and confidential by law (Title 13, United States Code, Sections 9, 141, 193, 214, and 221). This law protects your privacy and requires that you provide the information requested. Only persons sworn to protect Census Bureau data will have access to the information you provide—and no one else. Thank you.


Sincerely,



Charles Louis Kincannon
Director, U.S. Census Bureau

Enclosures

USCENSUSBUREAU
Helping You Make Informed Decisions

**DEPARTAMENTO DE COMERCIO DE LOS ESTADOS UNIDOS**
Administración de Economía y Estadísticas
Oficina del Censo de los EE.UU.
Washington, DC 20235-0001
OFICINA DEL DIRECTOR

29 de agosto de 2005

Estimado Residente:

La Constitución de los Estados Unidos decreta que se realice un censo en los Estados Unidos cada 10 años. En preparación para el Censo del 2010, la Oficina del Censo de los Estados Unidos está llevando a cabo el Censo Nacional de Prueba de 2005. El propósito de esta prueba es desarrollar nuevos métodos que harán el Censo del 2010 más fácil, más conveniente y menos costoso para los contribuyentes.

Hay dos maneras de proveer para el censo la información sobre todas las personas que viven en esta dirección. Por favor, **SÓLO** escoja una:

OPCIÓN 1 > Vaya a www.census.gov/census2005 para completar su cuestionario en línea. (En inglés solamente)

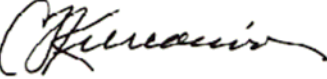
O

OPCIÓN 2 > Complete el cuestionario adjunto y devuélvalo por correo en el sobre prepagado que se incluye.

Su asistencia en ayudarnos a mejorar el censo es importante. Los resultados del Censo del 2010 se utilizarán para ayudar a cada comunidad obtener su parte justa de los fondos federales.

La ley (secciones 9, 141, 193, 214 y 221 del título 13 del Código de los Estados Unidos) requiere y protege la confidencialidad de sus respuestas. Esta ley protege su privacidad y requiere que usted provea la información que se solicita. Sólo las personas que han juramentado proteger los datos de la Oficina del Censo—y nadie más—tendrán acceso a la información que usted provee. Gracias.

Atentamente,



Charles Louis Kincannon
Director, Oficina del Censo de los Estados Unidos

Anejos

USCENSUSBUREAU
Helping You Make Informed Decisions

Form DC-1(E/S)
(6-2-2005)

040101 

2005 National Census Test Bilingual Form (Panel 20)

United States Census 2010 **2005 National Census Test**
 OMB No. 0607-0915; Approval Expires 06/30/2008

→ This is your official U.S. Census form. It is quick, easy, and your answers are protected by law. Please complete your form in the language most comfortable for you and return it today.

Start here

Before you answer Question 1, count the people living in this house, apartment, or mobile home using our guidelines.

Include these people:

- People who live here most of the time, even if they have somewhere else to live
- Roommates or boarders
- Children living here, including foster children
- People staying here on September 15, 2005 who have no other permanent place to stay

Do NOT include these people:

- College students who live away
- Armed Forces personnel who live away
- People who live or stay somewhere else most of the time
- People who, on September 15, 2005, are in:
 - a nursing home
 - a mental hospital
 - a correctional facility (example: jail)

1 How many people were living or staying in this house, apartment, or mobile home on September 15, 2005?
 Number of people -

2 Were there any additional people staying here September 15, 2005 that you did not include in Question 1? Mark all that apply.

Children, such as newborn babies or foster children

Relatives, such as adult children, cousins, or in-laws

Nonrelatives, such as roommates or live-in baby sitters

People staying here temporarily

No additional people

3 Is this house, apartment, or mobile home ...
 Mark ONE box.

Owned by you or someone in this household with a mortgage or loan?

Owned by you or someone in this household free and clear (without a mortgage or loan)?

Rented for cash rent?

Occupied without payment of cash rent?

4 What is your telephone number? We may call if we don't understand an answer.
 Area Code + Number

| | - | | - | | |

→ Begin with Person 1 on the next page.

Form DC-11ES (4-3-2005)

United States Census 2010 **Censo Nacional de Prueba de 2005**
 Núm. de OMB 0607-0915; Aprobado Hasta 06/30/2008

→ Este es su cuestionario oficial del Censo de los EE.UU. Es fácil y rápido de contestar y sus respuestas están protegidas por ley. Por favor, complete su cuestionario en el idioma que se sienta más cómodo y devuélvalo por correo hoy.

Comience Aquí

Antes de contestar la Pregunta 1, cuente a las personas que viven en esta casa, apartamento o casa móvil usando nuestras instrucciones.

Incluya a estas personas:

- Personas que se quedan aquí la mayor parte del tiempo, aunque tengan otro lugar donde vivir
- Compañeros de cuarto o inquilinos
- Niños que viven aquí, incluyendo hijos de crianza
- Personas que se quedaban aquí el 15 de septiembre de 2005 y que no tienen otro lugar permanente donde quedarse

NO incluya a estas personas:

- Estudiantes universitarios que viven fuera del hogar
- Personal de las Fuerzas Armadas que vive en otro lugar
- Personas que viven o se quedan en otro lugar la mayor parte del tiempo
- Personas que el 15 de septiembre de 2005 estaban en:
 - un hogar de convalecencia
 - un hospital para enfermos mentales
 - un centro de corrección (por ejemplo: una cárcel)

1 ¿Cuántas personas vivían o se quedaban en esta casa, apartamento o casa móvil el 15 de septiembre de 2005?
 Número de personas -

2 ¿Había personas adicionales quedándose aquí el 15 de septiembre de 2005 que usted no incluyó en la Pregunta 1? Marque todos los que apliquen.

Niños, tales como bebés recién nacidos o hijos de crianza

Parientes, tales como hijos adultos, primos o parientes políticos

Personas que no son parientes, tales como compañeros de cuarto o niñera que vive en el hogar

Personas que se quedan aquí temporalmente

No hay personas adicionales

3 ¿Es esta casa, apartamento o casa móvil ...
 Marque UNA casilla.

Propiedad suya o de alguien en este hogar con una hipoteca o préstamo?

Propiedad suya o de alguien en este hogar libre y sin deuda (sin una hipoteca o préstamo)?


Alquilada por pago en efectivo?

Ocupada sin pago de alquiler en efectivo?

4 ¿Cuál es su número de teléfono? Puede que lo llamemos si no entendemos una respuesta.
 Código de Área + Número

| | - | | - | | |

→ Comience con la Persona 1 en la próxima página.

040103  3

