The Effects Of Altering The Design Of Branching Instructions On Navigational Performance In Census 2000

FINAL REPORT

This research paper reports the results of research and analysis undertaken by the U.S. Census Bureau. It is part of a broad program, the Census 2000 Testing, Experimentation, and Evaluation (TXE) Program, designed to assess Census 2000 and to inform 2010 Census planning. Findings from the Census 2000 TXE Program reports are integrated into topic reports that provide context and background for broader interpretation of results.

Cleo Redline	Don Dillman	Aref Dajani	Mary Ann Scaggs
Statistical Research Division	Washington State University	Statistical Research Division	Statistical Research Division

USCENSUSBUREAU
Helping You Make Informed Decisions

CONTENTS

EXECUTIVE SUMMARY

It has been hypothesized and demonstrated that a number of languages (verbal, numeric, symbolic, and graphic) combine to affect respondents' perception and comprehension of visual information (Redline, C. and Dillman, D. 2002. "The Influence of Alternative Visual Designs on Respondents' Performance with Branching Instructions in Self-Administered Questionnaires." In Groves, R., Dillman, D., Eltinge, J., and Little, R. (eds.) *Survey Nonresponse*. New York: John Wiley and Sons, Inc.). Verbal language refers to the words on a questionnaire; numeric refers to the numbers; symbolic refers to symbols, such as arrows; and the graphic refers to the color, brightness, shape, and location of the information. Branching instructions were used as an initial test of this hypothesis because they yield objective measures of performance. A classroom experiment with college students, in which three of the languages (the verbal, symbolic, and graphic) were altered in two distinct ways (the *Prevention* and *Detection* methods) and tested against the Census 2000 method of branching, provided initial evidence in support of this hypothesis (Redline and Dillman 2002). In the Prevention method of branching, the response categories and check boxes were reversed. In the Detection method, a left-hand arrow came off the non-branching response option(s) and pointed to a paranthetical feedback phrase. This report documents an experiment conducted in Census 2000 in which the two branching instructions from the classroom experiment were revised, and two additional instructions were developed (*reverse printing* the instruction and substituting the words "*go to*" for "*skip to*") and tested against the Census 2000 version.

- The results of this experiment provide substantial evidence that manipulating the verbal, symbolic, and graphic languages of the branching instruction significantly influences whether the instructions are followed.

    Recommendation: Continue to gain expertise regarding the non-verbal (numeric, symbolic, and graphic) language of a questionnaire, for the non-verbal combines with the verbal to affect reading comprehension—that is, what respondents read (or do not read), the order in which they read it, and their consequent interpretation of what they read. This is an area of questionnaire design that is clearly emerging as critical to data quality and in need of further research.

- As hypothesized, respondents were no more apt to read the instruction 'go to' than 'skip to." The conclusion to be drawn from this is that no amount of rewording is going to help if the problem is respondents are not reading the information in the first place. In addition, respondents were more likely to mis-read the reverse printed instructions. Reverse-printed information may be more effective when a reader is actively searching for information as opposed to passively reading it.

    Recommendation: Before rewording information on a questionnaire, be certain that it is being read first. Cease using reverse printing anywhere on a

questionnaire, until more research can be done to explicate the conditions under which it is useful.

• The Detection Treatment significantly reduced commission errors (respondents answering questions they were instructed to skip) from 19.7 percent to 13.5 percent and omission errors (respondents not answering questions they were supposed to answer) from 5.0 to 4.0 percent.

> **Recommendation: Adopt the Detection method of branching in its present form, though further improvements may be possible. For example, it may be possible to combine the Prevention and Detection methods, and compare the results of this hybrid with either treatment alone. In addition, disentangle the contributing effects of the individual manipulations of the Detection treatment, since it is now clear that they are successful in combination.**

• Questionnaire design experts, respondents in cognitive interviews, and even a debriefing of respondents who participated in the classroom experiment revealed that many thought the arrows of the Detection method were confusing. However, the results of the Census experiment suggest otherwise (error rates actually decreased in the Detection Method). The conclusion to be drawn from this is: what people say and what people do are not necessarily one and the same.

> **Recommendation: Exercise care regarding what information we use to guide us when designing questionnaires. Relying on the verbal assessments of either respondents or expert questionnaire designers rather than performance measures from the experiments would have been misleading and detrimental in this case. Further research is necessary to determine when verbal reports are reliable and when they are not**.

• Wide variations existed in the error rates for individual questions. For example, on the Census 2000 Skip To Form, commission errors ranged from a low of 1.9 percent for the active duty question (question 20a) to a high of 79.2 percent for the age filter question (question 18). Also, significant differences existed between the error rates in the low (26.9 percent) and high (18.6 percent) coverage areas.

> **Recommendation: Analyze the potential reasons for these variations in future research, as well as the relationship between respondent characteristics and branching errors.**

• The overall pattern of commission error rates across treatments is similar between the classroom and the Census 2000 experiment. However, the absolute error rates within a treatment are either the same or higher in the census than the classroom.   This suggests

that respondents have a greater tendency to answer questions that do not apply to them under field conditions, despite the questions containing cues to the contrary and branching instructions. This is further evidence that respondents do not understand the questions or the underlying response task (that they need not answer every question). As a result, respondent burden is greater than necessary (on average, respondents are answering 20 percent more questions than they need to).

> **Recommendation: Further research is necessary to improve respondents' understanding of the questions and the response task and to reduce respondent burden.**

- The thesis of this research (that respondents extract meaning from more than just the verbal language of the questionnaire) has been borne out with branching instructions.

> **Recommendation: It now needs to be systematically extended to other areas of the questionnaire, like the questions themselves, and it is important to bear in mind that these issues extend to Web questionnaires as well.**

1. BACKGROUND

Information on a self-administered questionnaire can be decomposed into four language types: verbal, numeric, symbolic, and graphic (Redline and Dillman 2002).

- Verbal language – refers to the words.
- Numeric – refers to the numbers.
- Symbolic – refers to the check boxes, arrows, and other symbols on the questionnaire.
- Graphic – is the conduit by which all of the other languages are conveyed and includes the brightness, color, shape, and location of the information.

The major thesis of this program of research is that these languages combine to create meaning for respondents, and that with conventional branching instructions, three of these languages (the verbal, symbolic, and graphic) combine in such a way that respondents are often left unaware of the branching instruction.

One reason for this may be that, typically, these instructions are printed in the same font and point size as the rest of the text, making them difficult to detect (Foster 1979). In addition, Kahneman (1973) demonstrated that people's vision is sharp only within 2 degrees, which is equal to about 9 characters of text. Consequently, when a respondent is in the process of marking a check box, the branching instruction, which is usually located to the right of a response option, is likely to be outside of the respondent's view.   Also, this design does not take into consideration other strategies for reducing human error, like training respondents to prevent their errors in advance, or allowing them to detect errors afterwards (Norman 1990; Wickens 1992).

Thus, two new designs, the prevention and detection branching instructions, were developed, which manipulated the brightness, color, shape, and location (the graphic design) of the branching instructions, as well as incorporating prevention and detection strategies. Redline and Dillman (2002) offer a detailed description of these instructions, along with their depiction, which is briefly summarized here. In the prevention method, an instruction was placed before the question to remind respondents to pay attention to the branching instructions. The purpose of these reminders was to prevent mistakes before they were made. Also, the location of the response options and check boxes were reversed to bring the branching instruction into view and the branching instruction was made larger and bolder.

In the detection design, the branching instruction was made even bolder and larger to compensate for its poor location. Also, a left-hand arrow came off of the non-branching response options and pointed to the indented letter of the next question, which was followed by a parenthetical phrase. The purpose of these phrases was to allow respondents to detect and correct their mistakes after they were made. Consequently, both instructions attempted to make the verbal skip instruction more visible, but they differed in that the prevention

technique tried to remind people in advance that they may need to branch, whereas the detection technique gave them information afterwards, which allowed them to determine if they had branched correctly.

In a classroom experiment of 1,266 students, both designs were shown to decrease errors of commission (respondents answering questions they were instructed to skip) by more than half, from 20.3 percent on the control to 9.0 percent on the prevention form and 7.4 percent on the detection. However, errors of omission (respondents not answering questions they were instructed to answer) increased from 1.6 percent on the control to 3.3 percent on the prevention and 3.7 percent on the detection form (Redline and Dillman 2002) . In addition to the experiment, 48 cognitive interviews were conducted with a broad mix of people (Dillman et al. 1999; Redline and Crowley 1999).

Although respondents were supposed to understand that the check box and branching instructions were connected because they were next to each other in the same white background in the prevention instruction, both the results of the classroom experiment and the cognitive interviews suggested that this did not reliably work. Consequently, a stronger visual connection, an arrow, was devised for the census experiment (refer to Figure 1d). The reminder instructions may have contributed to the problem as well. Therefore, the number of these were dramatically reduced by strategically placing them after a long series of questions without any branching instructions. In addition, respondents demonstrated difficulty understanding the 'training' instruction of the prevention instruction, so it was simplified in the census design.

The larger size of the detection branching instruction in the classroom experiment appeared to overly attract respondents' attention to it, so this instruction was decreased in size for the census experiment. Also, respondents had trouble when they came to a branching instruction at the bottom of a page because the left-hand arrow did not point to anything. Thus, the left-hand arrow was made to terminate into another verbal branching instruction at the bottom of a page in the census design (refer to Figure 1e).

In addition to revising the prevention and detection instructions, two new instructions were developed to test additional issues. A branching instruction was designed to test the hypothesis that simply changing the verbal language from "skip to" to "go to" without making the instruction more visible is unlikely to make a difference in respondents' performance (refer to Figure 1b). Respondents in the cognitive interviews suggested that "go to" was clearer than "skip to." However, the purpose of the "go to" instruction was to demonstrate that rewording information may not prove to be the answer if the underlying problem is that respondents are not reading information in the first place. This is not to say, however, that rewording will not prove useful once respondents read the information. Rewording may very well improve comprehension, and consequently, performance further once respondents read the information, which explains our use of the words 'go to' in all of the experimental versions of the instruction.

Another branching instruction was designed to provide insight into whether printing branching instructions in reverse print is a good practice or not (Figure 1c). Normal print is the black lettering on the yellow background typical of most information on the census questionnaires. Reverse print is yellow lettering on a black background. There are arguments both for and against using reverse printing. On the one hand, it is plausible that the high contrast of a reverse-printed branching instruction and the fact that it is made visually dissimilar from the other information on the questionnaire could attract respondents' attention (Foster 1979) . On the other hand, typographical studies warn against using reverse print because it is difficult to read (Hartley, 1981; Wallschlaeger and Busic-Snyder, 1992). Also, since most of what respondents generally read is black, they may come to expect information to be printed in black. As a result, they may pay less attention to the occasional reverse-printed instruction.

2. METHODOLOGY

These ideas were tested in an experiment in Census 2000 using the long form. Five versions of the long form were developed, each employing a different treatment of the branching instruction. A sample of approximately 25,000 addresses was selected to receive one of the five treatments, with approximately 5,000 addresses independently selected per treatment. This number was distributed equally between so-called high coverage areas (2,500 per treatment), which are expected to have low proportion of minorities and renters, and low coverage areas (2,500 per treatment), which are expected to have a high proportion of minorities and renters.

Addresses on the Decennial Master Address File in the mailout/mailback areas of the country at the time sample selection took place served as the universe for sample selection. Consequently, addresses in non-mailback areas of the country (which can be characterized as highly rural areas of the country, where the forms need to be dropped off and picked back up by interviewers) were excluded from sample. Also, addresses that were added later as a result of coverage improvement operations were not included because these addresses were not available at the time of sample selection. Also, addresses that were in the Accuracy and Coverage Evaluation were excluded from sample so as not to overburden these households. A systematic sample by state, stratum (the high coverage and low coverage areas), and treatment was selected (Woltman, 1999). Analysis, then, was limited to those from this universe that mailed back their form.

The five treatments were the Census 2000 Skip To Instruction, Go To Instruction, (Go To) Reverse Print Instruction, (Go To) Prevention Instruction and the (Go To) Detection Instruction.

2.1 Treatments

2.1.1 The Census 2000 Skip To Instruction

Shown in Figure 1a, this instruction was used in the classroom experiment, and is exactly the same as the instruction used on the Census 2000 long form.

2.1.2 The Go To Instruction

Shown in Figure 1b, this instruction is like the Census 2000 instruction in all respects, except that the words "skip to" have been changed to "go to."

2.1.3 The (Go To) Reverse Print Instruction

Shown in Figure 1c, this instruction is like the Go To instruction, except that the words 'Go to" have been changed from normal print (black lettering on a yellow background) to reverse print (yellow lettering on a black background).

2.1.4 The (Go To) Prevention Instruction

Shown in Figure 1d, this is a modification of the prevention branching instruction from the classroom experiment, with "skip to" changed to "go to." A bold arrow was placed between the check box and the branching instruction to make the connection between these two pieces of information clearer. Also, the number of "attention" instructions was dramatically reduced. And finally, the language of the 'training instruction' was simplified from what it had been in the classroom experiment.

2.1.5 The (Go To) Detection Instruction

Shown in Figure 1e, this is a modification of the detection instruction from the classroom experiment, with "skip to" changed to "go to." The size of the branching instruction was decreased slightly from what it had been in the classroom experiment, and a left-hand arrow that terminated into a verbal branching instruction at the bottom of pages was added.

2.2 Implementation Procedures

The questionnaires in this experiment received *very nearly* the same implementation procedure as other questionnaires in Census 2000. The questionnaires were mailed out according to the Census 2000 schedule, with every sampled address mailed an advance letter, a questionnaire, and a follow-up postcard. One difference between the experimental and the census procedures, however, was that the experimental questionnaires were mailed back to the National Processing Center (in Jeffersonville, Indiana) rather than the

geographically desingated processing office. Consequently, the color of the return envelope was changed from white to yellow to facilitate its reaching the National Processing Office during the onslaught of census mail.

2.3 Analytic Procedures

2.3.1 Calculating Mail Response Rates

Households that returned duplicate forms were excluded from the calculation of the mail response rate (2 cases), as were households which did not return a form, but which were subsequently labeled as undeliverable as addressed in the mailout file (2834 cases). It was assumed in the latter case that the household was correctly labeled as non-existent or vacant. However, households that were identified as undeliverable as addressed in the mailout file, but which returned a questionnaire were included in the calculation. It was assumed in this case that the household was mistakenly labeled in the mailout file.

Non-response, then, was defined as any remaining household in the mailout universe which did not return a form, or returned a blank form. Blank forms were defined as having less than two answers for the first two persons. Response was defined as households from which a non-blank form from the mailout universe was received. The aggregate total for all responses (R) and for all nonresponses (NR) was established and then the total number of responses was divided by the total number of responses plus the total number of non-responses to yield the mail response rate (MRR), as shown

here: $MRR = \dfrac{R}{R + NR}$

2.3.2 Calculating Error Rates

To control for differences in the number of questions that respondents answered, this analysis was limited to the questions for Person 1. Branching error rates were calculated for questions that had branching instructions (because only their designs differed between form types) and those questions that had valid responses (because only then was it evident whether a respondent should branch or not).

An opportunity to make an error of commission occurred when a respondent selected a response with a branching instruction associated with it. An error of commission occurred if the respondent answered a question other than they were instructed to answer. An opportunity to make an error of omission occurred when a respondent selected a response that did not have a branching instruction associated with it. An error of omission occurred if the following question was left blank. It is also possible that lack of a response to a follow-up question may be due to a refusal by the

5

respondent to answer the question. However, it was assumed that refusals were equal across the treatment groups, and therefore, would not affect the conclusions drawn from the experiment. For any checked answer, either a commission opportunity or an omission opportunity can exist, but not both. Commission and omission opportunities, errors, and rates were calculated by respondent, by question within a treatment, across all questions, and across all treatments.

2.3.3 Significance Testing

Households were sampled at different rates within two geographic strata: high coverage areas and low coverage areas. Branching error rates were calculated by dividing the number of branching errors by the number of branching opportunities, where each of the two quantities is random. To compare rates across strata or treatments, standard errors were calculated using the statistical replication method of the stratified jackknife. As each household can have a variable number of branching opportunities and errors, clusters were incorporated into the variance estimation at the household level. Operationally, the stratified jackknife dropped one household at a time to calculate variance estimates. Statistical significance testing was conducted on pairs of strata or treatments using a t-test that incorporates the covariance between the branching opportunities and branching errors in the calculation. The normal approximation to the t distribution was used to calculate p-values and establish statistical significance. A Bonferroni adjustment was used to account for the multiple comparisons between treatments.

3. RESULTS

3.1 Response Rates

Response rates for the five treatment groups varied significantly between the high and low coverage areas, averaging 66.7 percent for the former and only 48.6 percent for the latter (Table 1). This difference of about 18 percentage points is not surprising inasmuch as the strata differ significantly with regard to the characteristics of residents. Respondents that mailed back their forms from the low coverage areas are about six times as likely as high coverage respondents to be nonwhite (53.8 percent vs. 8.8 percent) and/or Hispanic (23.7 vs. 4.25 percent). In addition, they are three times as likely to speak another language at home (28.7 percent vs. 9.6 percent) and more than twice as likely never to have graduated from high school (29.4 percent vs. 13.1 percent). These characteristics have often been associated with lower response rates for mail surveys (Dillman, 2000).

However, differences in response rates across treatment groups tend to be small or non-existent. Comparisons of the Skip To and Go To Control groups, and all Go To treatment groups with one another did not result in any significant differences.

6

3.2 Commission Error Rates

Significant differences existed for most of the treatment comparisons of commission error rates. It can be seen in Table 2 that the average commission error rate for the 19 branching items contained in the Census 2000 Skip form (Treatment 1) was 19.7 percent. As expected, the rate was substantially higher in the low coverage areas (26.9 percent compared to 18.6 percent).   Commission error rates also varied substantially by item (not shown in tabular form) with the overall rates ranging from 1.9 percent for the active duty question to 79.2 percent for the age filter question on the census form.

The overall commission error rate for the Go To comparison (Treatment 2) was not statistically different from the Skip to Treatment.

Because all of the remaining treatment groups used the words "Go to," Treatment 2 becomes the control group for the remaining comparisons. Table 2 reveals that all three of the remaining treatment groups achieved significantly   lower commission error rates, 17.9 percent for Reverse Print, 14.7 percent for the Prevention group, and 13.5 percent for the Detection group. The Detection Method, which had the lowest rate, was significantly lower than each of the other treatment groups, with two exceptions, the comparison with the Prevention group in the high coverage areas and at the national level. The Detection Method lowered the commission rate by about one-third for the strata as well as overall.

3.3 Omission Error Rates

Shown in Table 3, a somewhat different pattern emerges for the omission error rates. There were no significant differences between the standard Skip To form used in Census 2000 (Treatment 1), and the Go To (Treatment 2) version. In addition, the only form that significantly reduced the omission error rates was the Detection form, which did so across all groups.   In contrast, the omission error rates for the Reverse Print and Prevention Methods were significantly higher than the Go To Control form for all groups.   Thus, only the Detection form significantly reduced both the commission and omission error rates.

4. DISCUSSION

4.1 Response Rates

We hypothesized that changes in the branching instructions would have little effect on the response rates, and the results confirmed this (Table 1). When the data are parsed by high and low coverage area, the patterns are nearly identical, suggesting that for the most part, treatment and coverage area do not interact.

4.2 Commission Error Rates

There is only one change between the Skip To and Go To Treatments (the verbal instruction was changed from "Skip to" to "Go to"); therefore, the effect of this change is controlled for by design. However, several distinct changes were introduced into each of the next three treatments (Reverse Print, Prevention, and Detection), making it impossible to disentangle with certainty the effects of any single manipulation. Despite this limitation, however, the treatments build upon one another in such a way as to be highly suggestive regarding the individual manipulations and languages, as discussed below.

Figure 2 shows that, except for the Go To version, the commission error rate (for all areas) declines across the treatments. It was originally hypothesized that changing the instruction from "Skip to" to "Go to" would not affect the error rate because such a change does not address the underlying need to attract respondents' attention to the instruction first. The results support this hypothesis.

The fact that the commission error rate decreases across the Go To, Reverse Print, Prevention and Detection Treatments suggests that the changes made from one design to the next improved respondents' perception and comprehension of the instruction. We hypothesized that the Prevention and Detection Treatments would have this effect. However, the literature was ambiguous concerning what to expect from using reverse print. Although the Reverse-Printed Treatment led to a reduction in the commission error rates, this reduction was mediocre in comparison to the Prevention and Detection Treatments. This finding implies that there are probably better ways to attract respondents' attention to information on a questionnaire than using reverse print. It may be that respondents get used to reading information in a particular figure-ground (in this case, black against yellow). As a result, they come to expect that the information they should pay attention to will be black against yellow too. When the instruction is reverse printed, it may look so different that respondents have a greater tendency to disregard it. According to Gestalt psychologists a number of perceptual principles guide our understanding and interpretation of visual information. The Gestalt Grouping Law of Similarity states that we tend to see similar information as belonging together. Thus, reverse printing information may be an example of not using the visual element of color, or the Gestalt Grouping Law of Similarity, in a beneficial way (Wallschlaeger and Busic-Snyder, 1992).

The improvement in performance between both the Prevention and Detection Treatments over the other treatments is likely due to making respondents more aware of the branching instruction, i.e., making it more visible. However, the added improvement the Detection Treatment displayed over the Prevention Treatment in the low coverage area may be due to the feedback mechanism. It appears that this mechanism may have effectively helped these particular respondents to self-correct their mistakes.

4.3 Omission Error Rates

Figure 2 shows that the errors of omission decreased for the Detection Treatment in the census experiment, but increased for every other method. It would seem that unlike the Reverse Print and Prevention Treatments, which were intent on making the branching instruction more visible, the Detection Treatment successfully grouped information in the census experiment so that respondents were less likely to erroneously associate a branching instruction with a non-branching response option. It would seem that the left-hand arrow served the purpose intended here— to lead respondents' attention away from the branching instruction when they chose a response option without a branching instruction.

The implication of this finding is that grouping information correctly may be tantamount to respondents reading and understanding it correctly. If so, this is an example of the power of the Gestalt Grouping Law of Proximity (Wallschlaeger and Busic-Snyder, 1992), that is, of manipulating the visual element of location.

These results suggest that at the same time strong steps are taken to visually associate (or group) the check box(es) with the branching instruction(s), counter steps must also be taken to clearly dissociate the branching instruction from the other response options (i.e., not allow them to be seen as grouped together). It would seem that the Detection Treatment accomplished this balancing (grouping) act best.

In addition, the feedback mechanism may have worked better in the census experiment than the classroom experiment because it was simpler. In the census experiment it was almost always "(If Yes)" or "(If No)," whereas in the classroom experiment it tended to be a more complicated phrase, like "(If basketball, wrestling, or sent here from an earlier question)." For example, the errors of omission soared to 17.62 percent in the case of the citizen question in the Detection Treatment of the census experiment, whereas they averaged 4.05 percent across the rest of the questions using this same method. The feedback mechanism for the citizen question was "(If born outside or not a citizen of the United States)" whereas it tended to be "(If Yes)" or "(If No) for the rest of the questions.

4.4 Questionnaire Design Guidance

A conclusion to be drawn from this program of research is that we must be careful what information we use to guide us when designing questionnaires. Early on, when questionnaire design experts perused the questionnaires they had negative reactions to the use of the arrows in the Detection Treatment, claiming that the arrows looked confusing. Respondents in cognitive interviews often said the same (Dillman et al. 1999), and even a debriefing of respondents who participated in the classroom experiment revealed that some felt the arrows were confusing (Redline et al. 1999). These findings suggest that expert reviewers and respondents perceive themselves as being confused, when the outcome suggests otherwise. This may be occurring for one of two reasons. Preattentive processing

involves the automatic registration of features at global or holistic level, whereas attentive processing requires a detailed analysis of a field (Jenkins and Dillman 1997). Quickly surveying the questionnaire (that is, using preattentive processing rather than attentive processing) may lead to a perception of confusion when actually filling it out proves otherwise. And the other possibility is that respondents perceive themselves as confused because the feedback mechanism is working—it is slowing them down and helping them to correct their mistakes. Therefore, they perceive themselves as being lost and confused, when in reality their performance is improved.

Finally, it remains to be seen if the Prevention and Detection Treatments would benefit from being melded into one. One possibility is to maintain the reverse order of the check boxes/response options from the Prevention Treatment (because theoretically this should make the branching instruction the most visible), but counter balance this with the left-hand arrow and feedback mechanism, or something akin to this mechanism (because empirically this mechanism looks to be working) and compare the results of this hybrid to either treatment alone to see if it produces the greatest reduction in the errors of commission and omission.

4.5 Question and Respondent Effects

Redline and Dillman (2002) have proposed that respondents make mistakes navigating through a questionnaire both as a result of the characteristics of the questions they are answering (for example, questions that fall at the bottom of the page may lead to greater errors than those located elsewhere) and as a result of respondents' characteristics (for example, respondents with less education may make more mistakes). Differences in the error rates across the different versions of the branching instructions, across the different question types (an example of which was provided earlier with the citizenship question), and between respondent types (that is, between the coverage areas) provide evidence in support of both propositions. The conclusion to be drawn here is that this is a highly complex system under investigation, the effects of which are clearly going to take time to explore and explain well.

4.6 Effects of Cues from Follow-up Questions

The classroom experiment controlled for the effects of the wording of the questions so that respondents could get no cues from the questions themselves whether they should be answering them. However, in the census, the questions were dependent. So, for example, one of the questions asked respondents if they had any of their own grandchildren under the age of 18 living in their house or apartment, and if they did, then they were asked a follow-up question concerning whether they were responsible for these grandchildren. It seemed reasonable to expect that respondents would be able to figure out if a follow-up question applied to them in the census, not from reading the branching instruction, but from reading the content of the follow-up question and that the error rates would be lower in the census as a result.

However, a surprising finding to come out of this research is that the error rates are either the same or higher in the census than the classroom. This suggests that nationally representative respondents have a tendency to answer questions that do not apply to them, despite the fact that:

- the screener questions contain branching instructions, which clearly tell respondents to branch over the follow-on questions.
- the follow-on questions contain contrary cues, which ought to keep respondents from answering them.
- survey methodologists perceive the questions as logically connected, and therefore, think that respondents will too.

Therefore, not only may it be a good idea to improve upon the branching instruction, but also it may pay to recognize that respondents do not necessarily understand the basic logic of the questionnaire and the questions themselves, and that any steps taken to improve their understanding of this may help to reduce navigational errors as well.

REFERENCES

Dillman, D. 2000. *Mail and Internet Surveys: The Tailored Design Method.* New York: John Wiley & Sons.

Dillman, D., Carley-Baxter, L., and Jackson, A. 1999. "Skip Pattern Compliance in Three Test Forms: A Theoretical and Empirical Evaluation." *SESRC Technical Report #99-01* Social and Economic Sciences Research Center. Pullman: Washington State University.

Foster, J. 1979. "The Use of Visual Cues in Text." *Processing of Visible Language*, v.1. pgs. 189-201.

Hartley, J. 1981. "Eighty ways of Improving Instructional Text." *IEEE Transactions of Professional Communication*, 24, 1, 17-27.

Jenkins, C. and Dillman, D. 1997. "Towards A Theory of Self-Administered Questionnaire Design." In L. Lyberg, P. Biemer, M. Collins, L. Decker, E. DeLeeuw, C. Dippo, N. Schwarz and D.Trewin (eds), *Survey Measurement and Process Quality*. NewYork: Wiley-Interscience.

Kahneman, D. 1973. *Attention and Effort.* New Jersey: Prentice Hall.

Norman, D. 1992. *The Design of Everyday Things.* New York: Currency Doubleday.

Redline, C. and Dillman, D. 2002. "The Influence of Alternative Visual Designs on Respondents' Performance with Branching Instructions in Self-Administered Questionnaires." In Groves, R., Dillman, D., Eltinge, J., and Little, R. (eds.) *Survey Nonresponse*. New York: John Wiley and Sons, Inc.

Redline, C., Dillman, D., Smiley, R., Carley-Baxter, L. Jackson, A. 1999. "Making Visible the Invisible: An Experiment With Skip Instructions on Paper Questionnaires." *Proceedings of the Section on Survey Methods Research*, American Statistical Association.

Redline, C. and Crowley, M. 1999. Unpublished Data. Washington, D.C.: U.S. Bureau of the Census.

Wallschlaeger, C. and Busic-Snyder, C. 1992. *Basic Visual Concepts and Principles*. Dubugue, Iowa: Wm.C.Brown Publishers.

Wickens, C. 1992. *Engineering Psychology and Human Performance.* Second Edition. HarpersCollins Publishers, Inc.

Woltman, H. 1999. Sampling Specifications for the Research and Experimentation Program in K. Shaw, *Program Master Plan for the Census 2000 Alternative Questionnaire Experiment (AQE2000).* U. S. Bureau of the Census, Planning Research Evaluation Division, December 22, 1999.

Illustration of the five branching instruction treatments. (Figure 1)

a. Census 2000 Skip To Instruction	**30** **a. LAST YEAR, 1999, did this person work at a job or business at any time?** ☐ Yes ☐ No → *Skip to 31* **b. How many weeks did this person work in 1999?**
b. Go To Instruction	**30** **a. LAST YEAR, 1999, did this person work at a job or business at any time?** ☐ Yes ☐ No → *Go to 31* **b. How many weeks did this person work in 1999?**
c. (Go To) Reverse Print Instruction	**30** **a. LAST YEAR, 1999, did this person work at a job or business at any time?** ☐ Yes ☐ No → **Go to 31** **b. How many weeks did this person work in 1999?**
d. (Go To) Prevention Instruction	**30** **Attention: Remember to check for a "Go to" instruction after you answer the question below.** **a. LAST YEAR, 1999, did this person work at a job or business at any time?** Yes ☐ No ☐ → **Go to 31** **b. How many weeks did this person work in 1999?**
e. (Go To) Detection Instruction	**30** **a. LAST YEAR, 1999, did this person work at a job or business at any time?** ☐ Yes ☐ No → **Go to 31** **b. (If Yes) How many weeks did this person work in 1999?** *Count paid vacation, paid sick leave, and military*

Response Rates By Treatment (Table 1)

Treatment	Weighted National Total (All Areas)	Weighted N (in 000s)	High Coverage Areas	Unweighted N	Low Coverage Areas	Unweighted N
1. Census 2000 Skip To	63.9%	(12,630)	67.5%	(2,377)	48.5%	(2,321)
2. Go To Control	64.3	(12,520)	67.8	(2,355)	49.5	(2,307)
3. (Go To) Reverse Print	61.8	(12,650)	64.9	(2,387)	48.1	(2,304)
4. (Go To) Prevention	63.1	(12,540)	66.7	(2,365)	47.6	(2,288)
5. (Go To) Detection	63.3	(12,660)	66.4	(2,388)	49.3	(2,305)
Statistical Comparison						
1 vs. 2	n.s.		n.s.		n.s.	
2 vs. 3	n.s		n.s.		n.s.	
2 vs. 4	n.s.		n.s.		n.s.	
2 vs. 5	n.s.		n.s.		n.s.	
3 vs. 4	n.s.		n.s.		n.s.	
3 vs. 5	n.s.		n.s.		n.s.	
4 vs. 5	n.s.		n.s.		n.s.	

Commission Error Rates For All Census Long-Form Items With Branching Instructions (Table 2)

Instruction Treatment	Weighted National Total (All Areas)	Weighted Commission Opportunities (in 000s)	High Coverage Areas	Unweighted Commission Opportunities	Low Coverage Areas	Unweighted Commission Opportunities
1. Census 2000 Skip To	19.7%	(62,168)	18.6%	(2,302)	26.9%	(2,296)
2. Go To Control	20.8	(62,947)	20.0	(2,512)	25.4	(2,207)
3. (Go To) Reverse Print	17.9	(60,803)	16.7	(2,017)	24.9	(2,116)
4. (Go To) Prevention	14.7	(62,146)	13.6	(1,576)	21.7	(1,546)
5. (Go To) Detection	13.5	(61,046)	12.7	(1,653)	18.6	(1,830)
Statistical Comparison						
1 vs. 2	n.s.		n.s.		n.s.	
2 vs. 3	$p < .01$		$p < .01$		n.s.	
2 vs. 4	$p < .01$		$p < .01$		$p < .01$	
2 vs. 5	$p < .01$		$p < .01$		$p < .01$	
3 vs. 4	$p < .01$		$p < .01$		$p < .01$	
3 vs. 5	$p < .01$		$p < .01$		$p < .01$	
4 vs. 5	n.s.		n.s.		$p < .01$	

Omission Error Rates For All Census Long-Form Items With Branching Instructions (Table 3)

Instruction Treatment	Weighted National Total (All Areas)	Weighted Omission Opportunites (in 000s)	High Coverage Areas	Unweighted Omission Opportunities	Low Coverage Areas	Unweighted Omission Opportunities
1. Census 2000 Skip To	5.0%	(40,079)	4.8%	(1,649)	6.5%	(363)
2. Go To Control	5.4	(40,443)	5.2	(1,805)	6.3	(369)
3. (Go To) Reverse Print	7.6	(39,248)	7.3	(2,483)	9.1	(487)
4. (Go To) Prevention	7.0	(39,451)	6.7	(1,265)	9.4	(334)
5. (Go To) Detection	4.0	(39,044)	3.7	(2,241)	6.2	(510)
Statistical Comparison						
1 vs. 2	n.s.		n.s.		n.s.	
2 vs. 3	p < .01		p < .01		p < .01	
2 vs. 4	p < .01		p < .01		p < .01	
2 vs. 5	p < .01		p < .01		n.s.	
3 vs. 4	n.s.		n.s.		n.s.	
3 vs. 5	p < .01		p < .01		p < .01	
4 vs. 5	p < .01		p < .01		p < .01	

16

Error Rates by Panel. (Figure 2)