MAKING VISIBLE THE INVISIBLE: AN EXPERIMIENT WITH SKIP INSTRUCTIONS ON PAPER QUESTIONNAIRES

Cleo Redline, U.S. Bureau of the Census, Don Dillman, Washington State University, Richard Smiley, U.S. Bureau of the Census, Lisa Carley-Baxter and Arrick Jackson, Washington State University
Cleo Redline, U.S.Bureau of the Census, Room 3125—4, Washington, DC 20233

Keywords: Skip Instructions, Errors of Commission, Errors of Omission, Item Non-Response

Introduction

Research has shown that questionnaires with skip instructions produce higher item nonresponse than those without (Featherston and Moy, 1990: Messmer and Seymour, 1982). In addition, the failure to process skip instructions correctly leads to inconsistent responses between questions (e.g., Gower and Dibbs, 1989; Zuckerberg and Hess, 1996).

Two kinds of errors can be made as a result of skip instructions. An error of commission occurs when a respondent is instructed to skip over the following question (or questions), but instead answers it. An error of omission occurs when a respondent is supposed to answer the next question, but instead skips it. Both kinds of errors happen with substantial frequencies. For example, in an analysis of Census questionnaires, Raglin (1997) found that nearly 55 percent of the respondents who responded that they were working answered a series of questions about their non-work status (an error of commission). Conversely, 24 percent of the respondents who said that they were not working did not respond to the non-work questions (an error of omission).

A recently developed theory of self-administered questionnaire design posed visual techniques for getting respondents to follow a prescribed navigational path through a questionnaire (Jenkins and Dillman, 1995, 1997). Although this theory led to suggestions for gaining compliance with skip instructions, the suggestions were not empirically tested. Our purpose in this paper is to report results from an experiment in which two quite different skip instruction designs derived from the aforementioned theory were tested against the method of providing skip instructions used in the 2000 Census. The experiment was designed to test whether both errors of commission and omission could be reduced.

Theoretical Background

Many researchers have noted the problem of item non-response in self-administered questionnaires (e.g., Ferber, 1966; Craig and McCann, 1978). Others have related the graphical design of a form with difficulties respondents have completing the form (Wright and Barnard, 1978; Wright, 1980). A logical hypothesis is that item non-response results in part from the failure to visually define a clear navigational path through the questionnaire. To successfully define such a path requires taking into account perceptual processes. Jenkins and Dillman (1997) argued that information on a questionnaire could be reduced to its constituent visual elements of *brightness and color, shape, and location* and that these elements could be measured and intentionally manipulated to affect respondents' perception of the information.

Several authors have noted the problem of skip pattern compliance. Turner et al. (1992) reported that the faulty execution of skip instructions in self-administered questionnaires occurred to varying degrees, depending on a skip instruction's format. They concluded that respondents had a greater tendency to see information to the right of an answer category if it was somehow made salient. Jenkins and Dillman (1995) suggested that this was probably because the instructions were off to the right of where respondents' eyes were naturally traveling. Raglin (1997) measured the impact of alternative skip patterns on the collection of employment data from a test census. The differences between the skip instructions Raglin evaluated, however, were largely due to manipulating the order and content of the questions containing skip instructions as opposed to comparing the same verbal content with changes in visual presentation.

Skip patterns may pose a problem for a number of reasons. Cognitive interviews suggest that respondents' schemas of the form-filling task may be simpler than the form's actual requirements. Respondents often think they are supposed to read a question, read the answer categories, mark a box, and then look for the next question. However, what is really expected of them is more complicated.

A second reason skip instructions may pose a problem is because respondents may focus their attention on what they see as either the primary reason for the questionnaire, or the most interesting--the questions and response categories. This may reduce their attention to what they view as the less important or less interesting "mechanical" aspects of the questionnaire, like skip instructions.

Finally, a series of questions that do not contain any skip instructions may give respondents the erroneous expectation that they are supposed to answer the questionnaire sequentially. Because perception is sensitive to expectations, this could easily lead to

respondents overlooking skip instructions in future questions.

A central theme in the above cited reasons for making skip instruction errors is that respondents are not aware of the skip instruction at the moment they make an error. Therefore, it is incumbent upon the questionnaire to bring the skip instruction to respondents' attention. However, the design of the standard skip instruction does just the opposite. The problem is that the verbal skip instruction is printed in the same font and point size as the rest of the text. This makes it difficult to detect as something important to attend to and as something different from the remaining text (Wallschlaegar and Busic-Snyder, 1992).

In addition, the location of the skip instruction in the standard format is not conducive to its being perceived. Kahneman (1973) demonstrated that people's vision is sharp only within 2 degrees, which is equal to about 9 characters of text. This is known as the foveal view. However, when a respondent is in the process of marking a check box, the skip instruction, which is located to the right of a response option, would be located outside of the respondent's foveal view.

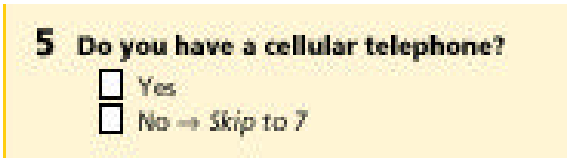
Finally, the present skip instruction design does not take into consideration that to err is human. Therefore, we need to understand human errors and ways of dealing with human errors. Besides making the necessary information visible, Norman (1990) suggests two additional strategies for reducing errors. The first is the *prevention* of errors through the use of mental aids. According to Norman, the notes we write to ourselves to remind us of tasks we need to accomplish are examples of mental aids. Another method for preventing errors, which doesn't assume we have learned the information already, is to train people in their prevention (Wickens, 1992). Effectively, training works by altering people's schemas about events. An example of being trained to prevent errors is being required to take a driver's education class before being allowed to drive.

The second strategy Norman suggests for reducing errors is to allow the user *to detect and correct* errors once they have occurred. Feedback can be used to allow users to detect errors. According to Norman, an example of feedback that most of us use everyday is the sound of our own voices when we speak.

Methods

Skip Instruction Treatments

To correct the kinds of deficiencies identified above in the standard skip instruction, two alternative skip instructions were developed. One uses the "prevention" strategy discussed above, and the other, the "detection" strategy. Both attempt, albeit in different ways, to make the verbal skip instruction more visible.



Treatment 1—Control Method

Shown in Figure 1, this treatment uses the standard Census Bureau skip instruction method in which the check boxes are on the left, and the response options are placed to their right. An arrow and a verbal skip instruction are provided to the right of the response option with no change in size (10 points) from the rest of the text (10-point).

Figure 1. The Control Skip Instruction

Treatment 2—Prevention Method

In this method, shown in Figure 2, respondents are provided an explanation of the skip instruction phenomenon before the first question containing a skip instruction. Secondly, an instruction "Attention: Check for a skip instruction after you answer the question below." is placed before every question with a skip instruction to remind respondents to pay attention to the skip instructions. In addition, this skip instruction relies on three techniques to make the printed verbal instruction more visible. First, the position of the check boxes and response categories are reversed, which makes it possible to place the printed verbal skip instruction immediately beside the check box and presumably within the foveal view of respondents. Second, the font, which was 10-point Frutiger Normal Italic in the control skip instruction, is 12-point Frutiger Black. Third, the background of the verbal skip instruction is changed from mustard (Pantone 129) to white to increase the contrast between the bold lettering and the background. Therefore, we hypothesize that, taken together, the following five manipulations will decrease skip instruction errors in self-administered questionnaires:

1. **Inclusion of a training instruction;**
2. **Inclusion of reminder instructions;**
3. **Decreasing the distance between the check boxes and the verbal skip instruction;**
4. **A larger, bolder font; and**
5. **Black text against a high contrast white background.**

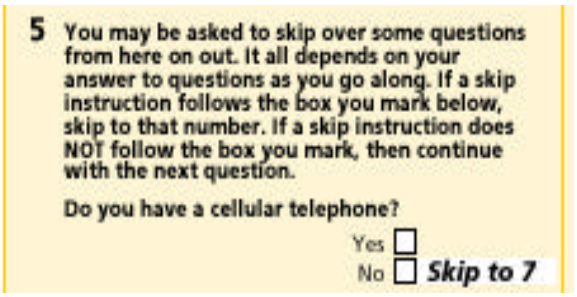


Figure 2. The Prevention Skip Instruction

Treatment 3—Detection Method

In this method, shown in Figure 3, the check boxes and the verbal skip instructions remain in their traditional locations. Since, as discussed earlier, this is not an ideal location for the skip instruction, the verbal skip instruction was made even bolder and larger (14-point Frutiger Black) to attract respondents' attention to this location.

Another difference between the detection version and the other versions is that the detection format is more explicit about the non-skip situation. In both the standard and the prevention formats, respondents are supposed to infer that in the absence of any explicit instructions to skip, they are to go to the next question. In the detection format, however, a bold arrow comes of the non-skip check boxes on the left-hand side and points to a parenthetical phrase at the beginning of the next question that succinctly repeats the meaning of the non-skip choices, e.g., "(If yes)." Together, the left-hand arrow and the parenthetical phrase are meant to provide feedback for respondents to self-correct their mistakes.

We, therefore, hypothesize the following manipulations will decrease skip instruction errors in self-administered questionnaires:

1. **A very large, bold font;**
2. **A left-hand arrow that connects non-skip check boxes with the next item; and**
3. **Inclusion of a feedback instruction at the beginning of the next item.**

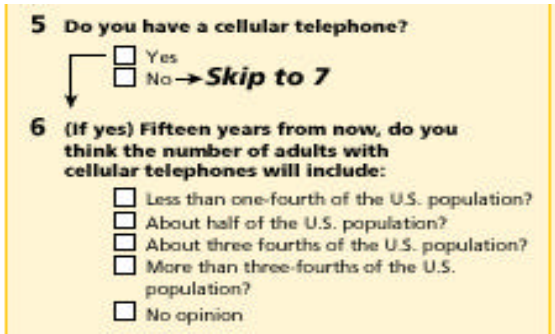


Figure 3. The Detection Skip Instruction

In sum, each of the hypotheses tests a combined strategy drawn from both visual and human performance theories. However, the strategies differ from one another. Largely, one aims at preventing errors and it reduces the distance between the check boxes and skip instructions, and the other retains the traditional distance and aims at detecting and correcting errors once they have been committed.

Test Questionnaire

In order to evaluate the three skip instruction formats, a four page 50-item questionnaire titled "Life Styles and Choices" was developed. Twenty-four of the questions contained skip instructions. One criteria for selecting and arranging the questions was that the respondent be given no clue from the content of surrounding questions as to whether a given question should or should not be answered. Three questionnaires were included in the test, each containing one of the treatment skip instructions.

Procedure

The questionnaire was administered to 34 classes of undergraduate students at Washington State University in October and November of 1998. The classes ranged in size from 9 to 192 students. Classes were selected from two campuses, Pullman, the main university campus located in Eastern Washington where students are primarily residential, and Vancouver, a commuter campus where students tend to be somewhat older and none live on campus. An attempt was made to vary the classes to obtain different kinds of students.

The questionnaires were systematically handed out in envelopes, with every third student receiving a different questionnaire. Students were instructed to return the questionnaire into the envelope immediately upon completion. The goal was to avoid a test mentality, whereby students who finished early might spend time providing answers to items they had missed.

Debriefing Questionnaire

After completing the test questionnaire, students were requested to fill out a one-page debriefing form, which asked them six questions about their experiences with the questionnaire and the skip instructions.

Sample Size

In total, 1266 students filled out the test and debriefing questionnaires: 422 students filled out the control form, 421 students filled out the prevention form, and 423 students filled out the detection form.

Evaluation Criteria

Compliance with the skip instructions was evaluated by comparing the *commission* and *omission* error percentages across forms. An error of commission occurs when a respondent is instructed to skip over the following question (or questions), but instead provides an answer to it. An error of omission occurs when a respondent is supposed to answer the following

question, but instead skips it. It is also possible that the lack of response to a question may be due to refusal by the respondent to answer the question. For this study, all such missing responses are assumed to be unintentional and due to an error of omission.

For each study question, the errors were calculated as follows: (1) Any respondent who provided no answer to the question (or who provided more than one answer to the question) was removed from the sample. (2) For those remaining, the answer to the study question was used to classify each respondent as supposed to skip or as not supposed to skip. (3) For those who were supposed to skip, responses to the questions to be skipped were used to determine whether or not a commission error had been committed. For each respondent, the total number of commission errors was computed. (4) For omission error candidates, response to the question following the study question was used to determine whether or not an omission error had been made. For each respondent, the total number of commission errors was computed.

Overall estimated commission and omission error rates were computed as weighted averages of the individual commission and omission error rates, with weights being the number of commission and omission error opportunities. Standard errors for the commission and omission error rates were computed using the VPLX variance estimation program using simple jackknife replication.

Results and Discussion

Test Questionnaire

Commission Error

The overall mean error percents shown in Table 1 suggest that both experimental forms outperformed the control form with respect to commission error, reducing the error by over 50 percent. The difference between the control form and either experimental form is significant at the alpha=.05 level. Differences between the two experimental forms were not significant.

Table 1: Overall Mean Percent of Commission Error by Form Type

	Control	Prevention	Detection
Error Percent	20.7%	9.0%	7.6%
Std. Error	(1.02)	(0.72)	(0.61)
No. of Errors	875	368	318
No. of Opportunities	4237	4102	4171
Sample Size	422	421	423

The distribution of the individual question errors across forms shown in Table 2 further supports these conclusions. The table reveals that while 16 of the detection form errors and 13 of the prevention form errors were less than 10%, only 3 of the control form errors were in that range. Conversely, while only 3 of the detection form errors and 3 of the prevention form errors were over 20%, 12 of the control form errors were in that range.

Table 2: Distribution of Commission Error Percents by Form Type

	Question Frequencies		
Error Percent	Control	Prevention	Detection
0 to 9.9%	3	13	16
10 to 19.9%	7	6	3
20 to 29.9%	6	2	2
30 to 39.9%	4	0	1
40 to 49.9%	1	0	0
50% or more	1	1	0

In summary, the data support the alternative hypotheses. The verbal and visual manipulations that comprised both the prevention and detection versions of the skip instructions successfully reduced the commission error percentages by more than half what they were in the control versions. However, it should be noted that the wide range of the error percents across questions suggests that other factors besides the format of the skip instruction were having an impact on level of compliance with the instruction. For instance, factors such as the skip instruction's placement on the page, the complexity of the question and response categories appear to have played a role. These factors will be investigated in future research.

Omission Error

The overall mean percents of omission errors shown in Table 3 suggest that the control form outperformed the experimental forms with respect to omission error. The average omission error on both experimental forms was more than double the error on the control form. The difference between the control form and both of the experimental forms is significant. The difference between the prevention form and detection form, however, is not significant.

Table 3: Overall Mean Percent of Omission Error by Form Type

	Control	Prevention	Detection
Error Percent	1.6%	3.3%	3.7%
Std. Error	(0.24)	(0.37)	(0.43)
No. of Errors	52	98	109
Opportunities	3217	3026	2954
Sample Size	422	421	423

Table 4 shows the distribution of the error percents across forms. The distribution appears to be similar across forms. The table also reveals that for most questions, the percent of omission error was much less than the percent of commission error: 16 of the control form omission errors, 12 of the detection form omission errors, and 13 of the prevention form omission errors were less than 3%.

Table 4: Distribution of Omission Error Percents by Form Type

	Question Frequencies		
Error Percent	Control	Prevention	Detection
0 to 2.9%	16	12	13
3 to 5.9%	2	4	3
6 to 8.9%	1	1	2
9 to 11.9%	0	1	1
12 to 14.9%	0	0	1
15% or more	1	2	0

In summary, the data do not support the alternative hypothesis. The verbal and visual manipulations that comprised both experimental skip instructions did not decrease the overall errors of omission. Instead, they increased them. However, according to Table 3, the majority of the questions had between a 0 to 2.9% error of omission across all forms. Therefore, the increase in the overall error of omission in both the experimental skip instructions was due to the presence of a few individual questions that weighted the means upwards in both. This suggests that if these few problems can be corrected, the overall mean on both the prevention and detection forms can be decreased.

In general it would appear that the more often respondents saw and executed the skip instruction correctly, the greater the likelihood that they saw and executed the skip instruction incorrectly, too. This seems to suggest that the feedback mechanism built into the detection format did not successfully prevent omission errors, at least it failed to do so in a number of questions. This suggests that we need to find a balance between making the skip instruction visible enough that respondents are likely to execute it correctly, but not so visible that they also execute it incorrectly.

Debriefing Questionnaire

Adding to our knowledge about the three skip instructions are respondents' answers to the debriefing question, as shown in Table 5. In terms of the debriefing questions, the prevention form seemed to outperform the others. Respondents were least likely to find instructions on the prevention form difficult to follow and least likely to find the skip instructions confusing and annoying. This finding seems to contradict the notion that the repeated use of the prevention instructions and the use of the word "Attention" in the prevention statements are bound to annoy respondents. Also, this finding seems to contradict the notion that the jagged left margins caused by right justifying information on the prevention form makes it more difficult to read, and again, is annoying. Although respondents were generally very capable of pointing out differences between the three forms in cognitive interviews, they were almost never aware that the check boxes and response options were reversed or that the left-hand margins were jagged. In general, these changes did not seem to bother them. Going back to our original theory, it may be that putting information in the order in which it is to be used, although not obvious to respondents, is satisfying nonetheless.

Table 5: Percent of Unfavorable Responses to Debriefing Questions by Form Type

	Percent of Responses		
Debriefing Question	Control *n=413*	Prevention *n=409*	Detection *n=420*
Found Questions Boring	18.9	17.4	15.5
Found Instructions Difficult to Follow	2.9	1.5 [3]	5.2
Not Concentrating on the Questions	10.1 [1,2]	6.1	4.8
Did not Like Something About the Form	27.2 [1,2]	20.2	19.9
Found Skip Instructions Confusing	4.8 [2]	3.2 [3]	9.3
Found Skip Instructions Annoying	21.3 [1]	14.4	18.1

[1] Control form percent is significantly different than Prevention form percent at alpha=.05 level

[2] Control form percent is significantly different than Detection form percent at alpha=.05 level

[3] Prevention form percent is significantly different than Detection form percent at alpha=.05 level

In contrast, respondents perceived the skip instruction on the detection form as more confusing than on the prevention form. Perhaps this was because of the arrows. Respondents seemed to be anything but neutral about the arrows in the cognitive interviews. They either loved or hated them. Or perhaps respondents found the detection system confusing because rather than preventing mistakes before they made them, it more or less forced them to correct their mistakes after the fact. Although it worked to decrease errors of commission, respondents may have perceived it as confusing.

With a couple of exceptions, the control form fell in between the prevention and detection forms. An interesting deviation from this occurred with more respondents saying that they didn't concentrate on the questions of the control form. It seems that the more likely respondents were to correctly execute the skip instructions, the more likely they were to say that they concentrated on the questions. Perhaps what respondents were really trying to say is that the experimental skip instructions required greater

concentration from them. This would be in line with our theoretical model in which executing the skip instructions correctly is an additional task over and above the other tasks of the questionnaire that requires greater cognitive resources of respondents.

Finally, respondents were most likely to say they did not like something about the control form and they were most likely to say that the control skip instructions were annoying. An interesting question, therefore, is: Will these results translate into differences in response rates? If so, the debriefing suggests that the control form may not fare well in comparison to the experimental forms in this regard either.

Conclusions

The experimental skip instructions show promise. They reduced the overall errors of commission by more than half, and although the errors of omission doubled, this increase appears to be due to a few questions only, and therefore, we think, correctable. Although there was no difference between the experimental forms in terms of error rates, respondents perceived the detection form as more confusing. Therefore, a mailout/mailback test needs to be conducted to determine what effect this may have on response rates. Also, a mailout/mailback test needs to be conducted to test the skip instructions under more natural conditions, with a representative sample of respondents and a real questionnaire.

We manipulated a large number of visual and verbal elements in the skip instructions, the effects of which we can only speculate about. It would be useful for future research to better control these manipulations, so we can begin to predict with accuracy the outcome of such manipulations.

Finally as we noted earlier, the wide range of the error percents across questions suggests that other factors besides the format of the skip instruction are having an impact. We intend to explore these factors in greater depth next.

References

Craig, C, and McCann, J.M. 1978. "Item nonresponse in mail surveys: extent and correlation." *Journal of Marketing Research* 15:285-89.

Featherston, F. and Moy, L. 1990. "Item Nonresponse in Mail Surveys." Paper presented at the International Conference of Measurement Errors in Surveys, Tucson, Arizona.

Ferber, R. 1966. "Item nonresponse in a consumer survey." *Public Opinion Quarterly*. 30:399-415.

Gower, A.R., and Dibbs, R. 1989. "Cognitive Research: Designing a Respondent Friendly Questionnaire for the 1991 Census." *Proceeding of the Bureau of the Census Fifth Annual Research Conference*, pp. 257-266.

Jenkins, C.R. and D.A. Dillman. 1995. "The Language of Self-Administered Questionnaires As Seen Through the Eyes of Respondents," *Seminar in New Directions in Statistical Methodology*, Statistical Policy Working Paper 23, Part 3 of 3, pp. 470-516.

Jenkins, Cleo R. and Don A. Dillman. 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*. New York: Wiley-Interscience.

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

Messmer, D. J. and D. T. Seymour. 1982. "The Effects of Branching on Item Nonresponse." *Public Opinion Quarterly*. Vol 46:270-277.

Norman, D.A. 1990. *The Design of Everyday Things*. New York: Currency Doubleday.

Raglin, D. 1997. "Using the 1996 National Content Survey to measure the Impact of Alternative Skip Patterns on the Collection of Employment Data." *American Statistical Association: Proceedings of the Section on Survey Research Methods*. pp. 342-347.

Turner, C.F., Lessler, J.T., George, B.J., Hubbard, M.L. and Witt, M.B. 1992. "Effects of Mode of Administration and Wording on Data Quality" in C.F. Turner, J.T. Lessler, and JC. Gfroerer (eds), *Survey Measurement of Drug Use Methodological Studies*, Washington, DC: National Institute of Drug Abuse, U.S. Department of Health and Human Services, pp. 221-243.

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

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

Wright, P. and Barnard, P. 1978. "Asking Multiple Questions about Several Items: The Design of Matrix Structures on Application Forms." *Applied Ergonomics*. 9,1,7-14.

Wright, P. 1980. "Strategy and Tactics in the Design of Forms." *Visible Language*, XIV 2, pp. 151-193.

Zukerberg, A. and Hess, J. 1996. "Uncovering Adolescent Perceptions:Experiences Conducting Cognitive Interviews. *American Statistical Association: Proceedings of the Section on Survey Research Methods*.

Acknowledgements

The authors wish to thank Bob Fay for his invaluable assistance in the analyses of these data, Eleanor Gerber for her helpful comments in the preparation of this paper, and Debbie Bolton and Henry Woltman for their support of this research. This paper reports the results of research and analysis undertaken by Census Bureau and Washington State Staff. It has undergone a more limited review than official Census Bureau publications. This report is released to inform interested parties of research and to encourage discussion.