INFLUENCE OF TYPE OF QUESTION ON SKIP PATTERN COMPLIANCE IN SELF-ADMINISTERED QUESTIONNAIRES

Don A. Dillman, Washington State University; Cleo D. Redline, U.S. Bureau of the Census;
Lisa R. Carley-Baxter, Washington State University*
Don A. Dillman, Washington State University, PO Box 644014, Pullman, WA 99164-4014

Key Words: Skip instructions, Errors of commission, Errors of omission, Item nonresponse

A recent paper (Redline, Dillman, Smiley, Carley-Baxter, and Jackson, In Press) reports results from a test of two methods of visual design for improving compliance to skip instructions. One procedure, the *detection method*, relied on the provision of visual cues to respondents that were designed to help them detect if they had made a skip error. A second procedure, the *prevention method*, provided visual clues aimed at preventing navigational errors before they happened. When compared to a traditional (control) method of providing skip instructions, both of these methods significantly reduced errors of *commission* (the tendency not to skip when directed to do so). Specifically, the prevention method achieved an overall error rate of 9% and the detection method had an overall error rate of 7.6% compared to a 20.7% overall error rate for the control group. However, when compared to the control method, these two methods did not reduce *omission errors* (the tendency to skip over questions that were supposed to be answered). For omission errors, the control method had an error rate of only 1.6%, compared to a rate of 3.7% for the detection method and 3.3% for the prevention method.

Substantial differences for skip pattern compliance of both types were also reported for the 24 test questions. The percent of respondents making errors of commission ranged from 0 to 52%, and omission error rates ranged from 0 to 34%. And, the higher the error rate for a question on one form the higher it tended to be on another. The zero order correlations for commission errors were strongly positive, .66 for the prevention and detection forms, .82 for the prevention and control form, and .84 for the detection and control forms. The correlations for omission errors were more varied, but also positive, being .30 for the prevention and detection forms, .37 for the control and detection forms, and .90 for the prevention and control forms. These positive correlations raise the issue of whether the rates for skip-pattern compliance for either or both types of errors could be partly explained by the structural characteristics of the individual questions.

Our purpose in this paper is to describe eight distinguishing structural attributes of the test questions and examine through regression analyses their potential influence on respondent errors of commission and omission.

Development of Two Strategies for Improving Skip-Pattern Compliance

The theoretical background for development of the prevention and detection methods is reported in two companion papers (Redline, Dillman, Smiley, Carley-Baxter, and Jackson, In Press; Redline and Dillman, Forthcoming).

Treatment 1 – Control Method

This treatment uses a method of providing skip instructions 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 from the rest of the text (10-point). However, the verbal skip instruction is changed slightly in shape and brightness from the response options. This is the skip instruction method that will be used by the U.S. Census Bureau for the 2000 Census long form questionnaires, and illustrated by the first box of each example in Figure 1.

Treatment 2 – Prevention Method

In this method five techniques are combined to increase the prominence of the skip directions. They include:

1. **Educating respondents about skip instructions before the first question containing a skip instruction;**
2. **Reminding them to pay attention to the skip instructions before every question that contains a skip instruction;**
3. **Decreasing the distance between the check boxes and the verbal skip instruction by placing answer boxes to the right of each response choice;**

The authors wish to thank Dr. Rodney K. Baxter, Washington State University and Dr. Robert Fay, U.S. Bureau of the Census, for their assistance with the regression analysis procedures used in this paper.

4. **Printing the verbal skip instruction in a larger, bolder font; and**
5. **Placing the black text of the verbal skip instruction against a high contrast white background.**

This procedure is shown in the third box of each example of Figure 1.

Treatment 3– Detection Method

In this method, the check boxes remain in their traditional location on the left-hand side of a column and the verbal skip instructions remain in their traditional location to the right of the response options. Since this is not an ideal location for the skip instruction (Redline and Dillman, Forthcoming), the verbal skip instruction was made even bolder and larger to attract respondents' attention to this location. The three elements of this detection method, illustrated in the second example of each question in Figure 1, are as follows:

1. **Printing the verbal skip instruction in a very large, bold font;**
2. **A left-hand arrow that connects the check boxes of respondents who are not supposed to skip the next item; and**
3. **An instruction at the beginning of the next item that explains who should answer that item.**

Procedures

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. Three questionnaires were included in the test, each containing one of the treatment skip instructions. The test questionnaires were administered to 34 classes of undergraduate students at Washington State University in October and November of 1998. In total, 1,266 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.

Effects of Type of Question

Different types of questions place different cognitive demands on individuals, e.g., having fewer or more answer choices which must be compared in selecting an answer, the number of words that must be mentally processed, and the presence of different skip directions for different response categories. Eight attributes that differentiate questions from one another are identified here, each of which is hypothesized as increasing the likelihood that skip and/or continuation directions will not be followed correctly. Generally, these variables reflect the proposition that greater cognitive complexity increases competition for the respondent's attention, making it less likely that skip instructions will be seen, comprehended, and the direction they provide followed.

Characteristics Hypothesized to Affect Error Rates

Location at the bottom-of-the-page Questions at the the bottom-of-the-page (particularly those that occur before turning a page) interrupt the cognitive processing the respondent is engaged in while filling out the questionnaire and subsequently their attention may wander, thereby causing them not to comply with a skip instruction. Respondents are also likely to have a schema (a "collection of knowledge from past experience" (Eysneck, 1993:37)) about how to answer surveys and are used to answering every question which instinctively leads to answering questions at the top of the next column, regardless of whether they've been previously instructed to skip that question. This variable was dummy coded for the regression analysis, 0 if the question is not at the bottom-of-the-page and 1 if the question is at the bottom-of-the-page. Seven items have this characteristic.

First categories continue/Last categories skip Questions in which the first answer category(ies) direct(s) the respondent to continue to the next listed question and the last category(ies) direct(s) people to skip have an unusual navigational aspect. The respondent who chooses one of the first answers must visually pass over additional categories containing skip instructions. It seems plausible that some respondents may be attracted to these last visual aspects of the question they have just answered and mistakenly conclude that these skip instructions, located beside an answer choice they did not choose, also apply to them. An example of this situation is provided for each treatment in Figure 1, Item A. This variable was dummy coded for the regression analysis, 0 if the question does not have a first category continues/last category skips and 1 if it does have a first category continues/last category skips. Ten items have this characteristic.

Every answer skips Questions that require every answer to skip to a different question present an increased amount of information to respondents and are quite complex (Item B, Figure 1). Respondents may not be able to process all of this information or may simply see the wrong instruction (that is, one that is not associated with the answer category they chose, but rather one associated with an answer category above or below their answer choice). Conversely, the attention of their short-term memory may have been directed elsewhere (Neisser, 1967) thereby contributing to an error in following the skip instruction. This variable was dummy coded for the regression analysis, 0 if all the answer categories do not skip and 1 if they all do have a skip instruction. Four items have this characteristic.

Write-in answers Questions that require write-in answers also disrupt the normal cognitive processing that a respondent goes through when answering a question, and may conflict with a respondents' schema about how to answer survey questions (Item C, Figure 1). Respondents are used to checking a box that corresponds to their answer and then answering the next question. However, with write-in answers respondents are asked to write their answer in an answer box and then proceed. After completing this task, respondents may not see the skip instruction, or forget to look for such an instruction because their pattern of answering questions by just checking boxes has been interrupted. This variable was dummy coded for the regression analysis, 0 if the question is not a write-in answer and 1 if it is a write-in answer. Two items have this characteristic.

Alternating skips Questions with alternating skip instructions are quite complex. They require some answer categories to skip while others do not—however, they do so in a manner in which one category skips, the next doesn't, and a third does. Alternating skips may confuse the respondent because there is so much information to process. Or respondents may fail to recall the skip instruction because the attention of their short-term memory has been directed elsewhere (Neisser, 1967). This variable was dummy coded for the regression analysis, 0 if the question does not have alternating skip instructions and 1 if it does. Four items have this characteristic.

Distance between the answer box and skip instruction Questions with more distance between the answer box and the skip instructions may have higher error rates because people can only sharply see about 8-10 characters at one time (Kahneman, 1973). Therefore, as respondents are checking the answer box, the skip instruction may too far away to be in their direct visual field. If they can't see the skip instruction, then they can't comply with it. We used the median distance between the answer box and skip instruction (range 0 to 2.62) to divide answers into low distance (0 - .5 inches) or high distance (.51 inches or more) categories.

Number of answer categories Questions with a higher number of answer categories may cause the respondent some confusion due to the sheer number of choices. This confusion may be a part of the limitations of short-term or working memory (Miller, 1956). Questions with a higher number of response categories may place more demands on the storing of that information in short-term memory (Eysenck, 1993) thereby contributing to higher error rates. However, higher error rates on questions with a high number of answer categories may also be caused in part by respondents confusion about where they are supposed to go and/or looking at the wrong skip instruction. We used the median number of answer categories (range 2 to 11) to divide answers into low (0 up to 3) and high (4 or more).

Number of words Questions with a higher number of words in the text of the question may have higher error rates because of the limitations of short-term memory (Eysneck, 1993). We used the median number of words (range 5 to 62) in the question text to divide the answers into low (up to 19 words) or high (20 words or more) categories.

Operationalization of Dependent Variable

Any respondent who provided no answer to the question (or who provided more than one answer to the question) was removed from the sample. For those remaining in the sample, the answer to the study question was used to classify each respondent as *supposed to skip* or as *not supposed to skip.* For those who were supposed to skip, the presence or absence of response to the questions to be skipped was used to determine whether or not a *commission error* had been committed. The number of those who made a commission error was used as the numerator in the commission error percent calculation and the number of those who were supposed to skip was used as the denominator.

For *omission error* candidates, the presence or absence of a response to the question following the study question was used to determine whether or not an omission error had been made. The number of those who made an omission error was used as the numerator in the omission error percent calculation and the number of those who were not supposed to skip was used as the denominator.

Results

Bivariate Analyses

We hypothesized that each question characteristics will have higher commission and omission error rates than questions that do not have these characteristics. As shown in Figure 2, we found that for 6 of the 8 question characteristics the *commission error* rates were in the predicted direction. However, questions for which the first response category(ies) directed the respondent to skip while the last category(ies) directed them to continue and questions with a high number of words in them had significantly lower commission errors than questions that did not posses these characteristics, which is opposite to what we hypothesized. When comparing *omission error* rates, we found that questions at the bottom-of-the-page, questions that require a write-in answer, and questions with a left-arrow above, all had higher omission error rates than questions without those characteristics—however, none of the differences were significant.

Figure 1. Examples of selected control, detection, and prevention formats for selected types of questions.

A. First category(ies) direct(s) respondent to continue while last category(ies) direct(s) respondents to skip.

24 **Have you ever purchased the sound track of a movie?**
☐ Yes
☐ No → *Skip to 26*

24 **Have you ever purchased the sound track of a movie?**
☐ Yes
☐ No → ***Skip to 26***

24 **Attention: Check for a skip after you answer ... Have you ever purchased the sound track of a movie ?**
Yes ☐
No ☐ ***Skip to 26***

B. All respondents must skip over the next question. This type of question makes *omission errors* impossible.

35 **About how often do you attend religious services?**
☐ Once a week or more
☐ 2-3 times a month } *Skip to 39*
☐ Once a month
☐ 4-9 times a year } *Skip to 37*
☐ Once a year or less
☐ Not at all } *Skip to 38*

35 **(If one of last two categories) About how often do you attend religious services?**
☐ Once a week or more
☐ 2-3 times a month } ***Skip to 39***
☐ Once a month
☐ 4-9 times a year } ***Skip to 37***
☐ Once a year or less
☐ Not at all } ***Skip to 38***

35 **Attention: Check for a skip after you answer ... About how often do you attend religious services?**
Once a week or more ☐
2-3 times a month ☐ } ***Skip to 36***
Once a month ☐
4-9 times a year ☐ } ***Skip to 37***
Once a year or less ☐
Not at all ☐ } ***Skip to 38***

C. Write-in answers: a type of question which interrupts normal answer procedure of checking boxes.

13 **Who is your favorite sports star?**
→ *Skip to 17*
☐ Mark this box if you don't have one.

13 **(If basketball, wrestling, or sent here from an earlier question) Who is your favorite sports star?**
→ ***Skip to 17***
☐ Mark this box if you don't have one.

13 **Attention: Check for a skip after you answer ... Who is your favorite sports star?**
Skip to 17
Mark this box if you don't have one. ☐

Figure 2. Error rates for questions with and without specified characteristics.

	Commission			Omission		
Qstn. Char.	With	W/out	Prob	With	W/out	Prob
Bottom-of-page	15.1	11.0	.02	5.2	2.3	.21
All skip	16.2	11.3	.01	0.0	2.9	---
Write-in	34.1	11.7	.00	5.0	2.7	.50
Alternating skips	14.0	12.3	.49	1.2	3.3	.60
1st cat. cont/last skip	5.4	13.5	.04	5.8	2.4	.16
High distance	15.3	9.9	.00	1.7	3.8	.39
High # answer cat.	17.6	10.5	.00	1.6	3.4	.52
High # words in ques.	8.7	14.1	.01	2.3	3.2	.68

Regression Analysis

A regression analysis was run that included all eight question characteristics hypothesized to have an influence on the *commission error* rate. Separate models were calculated for each type of form (control, prevention, detection) and the combination of all forms. For both the overall model and the individual form models only one variable, write-in answer, was a significant predictor of commission error rates. The overall model that included all types of forms was significant with 35% of the variance in error rates being explained (Figure 3). For the detection form model 67% of the variance was explained as was 62% for the prevention form. The over-all model for the control form was not statistically significant.

Figure 3. Regression analysis of *commission errors* for all three forms.

Independent Variable	Unstandardized Regression Coefficient	Standard Error	Standardized Regression Coefficient
Bottom-of-the-Page	3.69	3.13	0.14
All Answers Skip	5.90	4.31	0.16
Write-In Answer	25.16***	4.82	0.56
Alternating Skips	1.25	4.19	0.04
# of Answer Categories	0.08	0.60	0.02
Last Category Skip	0.78	2.92	0.03
# of Words in Question	-0.13	0.12	-0.12
Intercept	11.81		
R-squared	.35***		
Model df	7		
Total df	70		

*p<.05, **p<.01, ***p<.001

Because of the powerful effects of the write-in answer, each of the models was recalculated deleting the write-in answer variable. This resulted in none of the variables being significant for either the combined or individual form models. Nor was the r-square for any of these models significant. We then conducted an analysis in which data for the control versus detection forms were combined because of their structural similarity (answer boxes to the left). A new variable, the distance between answer box and skip instructions, was added to this analysis along with a form (control versus detection) variable (Figure 4).

Three variables are significant in the control/detection form analysis, including write-in answer, distance from answer box to skip instruction, and the control verus detection form. The overall model is significant with 63% of the variance being explained by these three variables. This model was re-run with the addition of eight interaction terms (each question characteristic by detection form). In this model only write-in answer was significant. The overall variance explained was 65%, about the same as when no interaction terms were included.

Figure 4. Regression analysis of *commission errors* for control and detection forms only.

Independent Variable	Unstandardized Regression Coefficient	Standard Error	Standardized Regression Coefficient
Bottom-of-the-Page	4.48	3.31	0.16
All Answers Skip	8.50	4.44	0.22
Write-In Answer	26.19***	4.99	0.56
Alternating Skips	2.25	4.33	0.06
# of Answer Categories	-0.18	0.63	-0.04
Distance/answer box & skip	5.45*	2.63	0.24
Last Category Skips	0.99	3.19	0.04
# of Words in Question	-0.07	0.16	-0.05
Detection	-13.57***	2.63	0.52
Intercept	12.85		
R-squared	.63***		
Model df	9		
Total df	46		

*p<.05, **p<.01, ***p<.001

The analysis for *omission errors* exhibited quite different results. First, for the combined analysis of all forms, the overall model was significant with 25% of the variance being explained (Figure 5). One variable in the model, bottom-of-the-page location of the question, was a significant predictor of error rates. The individual analysis of the detection form resulted in 59% of the variance being explained, which was also significant. In this model, two of the individual variables were significant, the bottom-of-the-page and last category skips. The omission error individual analyses for the control and prevention forms resulted in no variables being significant and a similar lack of significance for each of the overall models.

Figure 5. Regression analysis of *omission errors* for all three forms.

Independent Variable	Unstandardized Regression Coefficient	Standard Error	Standardized Regression Coefficient
Bottom-of-the-Page	5.66**	1.73	0.42
Write-In Answer	4.29	2.44	0.22
Alternating Skips	0.11	2.14	0.01
# of Answer Categories	-0.60	0.31	-0.28
Last Category Skips	-2.23	1.49	-0.19
# of Words in Question	0.05	0.06	0.11
Intercept	4.07		
R-squared	.25**		
Model df	6		
Total df	61		

*p<.05, **p<.01, ***p<.001

An additional analysis was run which included all data from the control and detection forms. Limiting the analysis to these two forms allowed us to include an additional independent variable, distance between answer

box and skip instruction. In this model two variables were significant, bottom-of-the-page and distance between the answer box and skip instruction. The overall variance explained was 43%, which was statistically significant. At this point, interaction terms were added for each of the seven question characteristics. In this model, three of the original variables—bottom-of-the-page, distance between answer box and skip instruction, and last category skips—were significant (Figure 6). In addition, the interaction between last category skips and the detection form was significant. Addition of the interaction terms to the model resulted in a substantial increase in the amount of variance explained from 43% to 62% in the model.

Figure 6. Regression analysis of *omission errors* for control vs. detection forms with interactions.

Independent Variable	Unstandardized Regression Coefficient	Standard Error	Standardized Regression Coefficient
Bottom-of-the-Page	6.06**	2.01	0.59
Write-In Answer	1.45	2.80	0.10
Alternating Skips	-0.82	2.40	-0.07
# of Answer Categories	-0.15	0.35	-0.10
Distance/answer box & skip	-3.49**	1.29	-0.48
Last Category Skips	-5.26**	1.79	-0.60
# of Words in Question	-0.07	0.09	-0.15
Detection	-2.67	3.22	-0.30
Bottom*Detection	-1.83	2.79	-0.14
Write-In*Detection	-2.91	3.94	-0.14
Alt Skip*Detection	1.54	3.43	0.10
# Ans Cat*Detection	-0.27	0.52	-0.17
Distance*Detection	0.73	2.21	0.10
Last Cat Skips*Detection	7.45**	2.51	0.70
Words*Detection	0.07	0.12	0.17
Intercept	8.42		
R-squared	.62**		
Model df	15		
Total df	40		

*p<.05, **p<.01, ***p<.001

Finally, the control and prevention forms were combined for analysis. For these forms, bottom-of-the-page and last category skips were both significant as was the overall model in which 36% of the variance was explained. Addition of the interaction terms resulted in none of the variables or the overall model being significant.

Discussion and Conclusion

The difference in *commission error* rates is primarily influenced across all forms by the occurrence of write-in answers. Its significance is not surprising. When a respondent comes to a write-in answer, a quite different response behavior than that to which they have become accustomed is required. It also appears from the analysis of combined results from the control and detection forms that a greater distance between answer box and skip instructions plays some role in the occurrence of commission errors.

Something quite different appears to be happening with regard to the influence of question characteristic on the occurrence of *omission errors*. Location at the bottom-of-the-page appears to be the primary influence on whether omission errors occur. This is not surprising, inasmuch as a person must shift from the bottom of a column or page to the next column and this appears to make it easier to forget exactly where one is supposed to go. It was also found that whether the last of the category choices required a skip influenced omission error rates only on the detection form. This too is not surprising, and suggests a potential problem with the left-hand arrow used to guide respondents from their answer to the next listed question, when that question doesn't appear below the arrow. We suspect that the respondent's having seen that arrow as they perused the top categories was overridden by seeing the skip instruction for the last answer category.

These findings suggest that to lower *commission error* rates we need to focus on improving instructions for write-in answers. However, to improve *omission error* rates we need to focus on how to better connect people from the last question on a page to the first questions on the next page where the visual connection is necessarily broken. Another possible implication is that questionnaire designers should avoid placing questions at the bottom-of-the-page that happen to contain skip instructions.

References

Eysneck, M. W. 1993. *Principles of Cognitive Psychology*. Hove, UK: Lawrence Earlbaum Associates.

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

Miller, G. A. 1956. "The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information." *Psych. Review* 63:81-97.

Neisser, Ulric. 1967. *Cognitive Psychology*. New York: Appleton-Century-Crofts.

Redline, C. and D. A. Dillman. Forthcoming. "The influence of auxiliary, symbolic, numeric and verbal languages on navigational compliance in self-administered surveys." In Groves, R. et al., *Nonresponse error in surveys*. New York: Wiley Inter-Science.

Redline, C., D. A. Dillman, R. Smiley, L. Carley-Baxter, and A. Jackson. In Press. "Making Visible the Invisible: An Experiment with Skip Instructions on Paper Questionnaires." Proceedings of the Section on Survey Research Methods (American Association for Public Opinion Research). American Statistical Association, Alexandria, VA.