Eye-Movement Analysis: A New Tool for Evaluating the Design of Visually Administered Instruments (Paper and Web)

Cleo D. Redline, Bureau of the Census, and Christopher P. Lankford, University of Virginia
Cleo D. Redline, Bureau of the Census, Room 3215—4, Washington DC 20233

Keywords: Surveys; Self-administered questionnaire design; Branching instruction errors

A data collection instrument that a respondent self-completes through the visual channel, such as on paper or over the Web, is visually administered. Although insightful in many ways, traditional methods of evaluating questionnaires, such as cognitive interviewing, usability testing, and experimentation may be insufficient when it comes to evaluating the design of visually administered questionnaires because these methods cannot directly identify information respondents perceive or the precise order in which they observe the information (Redline et al 1998).

In this paper, we present the results of a study that was conducted to explore whether eye-movement analysis might prove a promising new tool for evaluating the design of visually administered questionnaires. Eye tracking hardware and software, which were originally developed in the Systems Engineering Department of the University of Virginia for use with computer monitors, were adapted to track the eye movements of respondents answering three versions of a paper questionnaire. These versions were chosen for study because differences in the design of their branching instructions were hypothesized to affect eye-movements, which in turn may affect the accuracy of following the branching instructions (Redline and Dillman Forthcoming).

Background

Eye-movement analysis has been used in other fields, most notably reading and scene perception, to study cognitive processing (e.g., Rayner 1992; Rayner 1983). However, survey design research grew out of the interviewer-administered realm, which has been primarily focused on respondents' comprehension of the spoken language of questionnaires. Therefore, the mechanism by which respondents perceive information presented on paper questionnaires or over the Web, the eyes and their movements, has not received much attention until recently. Other reasons for the lack of eye-movement research in the survey field are its cost and relative difficulty. As others have noted, eye-movement research requires specialized knowledge, equipment and expertise to operate the equipment. In addition, the data are time consuming and difficult to analyze (Ellis et al. 1998; Lohse 1996).

Paper questionnaires typically contain instructions to advance a respondent to a particular question as a result of their response to the current question. However, respondents often do not follow these instructions (e.g., Turner et al. 1992; Featherston and Moy 1990; Messmer and Seymour 1982). Redline and Dillman (In Press) propose that a number of languages (visual, symbolic, and verbal) combine to affect respondents' perception and comprehension of branching instructions, and consequently, the navigational path of a form. Evidence for this assertion comes from a pilot study with college students in which these languages were altered in two distinct ways and tested against the Census 2000 branching instruction. The new designs were shown to decrease errors of commission (respondents answering questions they were instructed to branch). However, errors of omission (respondents not answering questions they were instructed to answer) increased. In this paper, we attempt to answer the question: does eye-movement analysis shed additional light on respondents' processing of branching instructions?

Methods

Questionnaire

Eye-movement analysis was conducted with the four-page questionnaire developed by Redline and Dillman (In Press). The questionnaire asked 50 questions about life styles and choices. Twenty-three of the questions contained branching instructions. Three versions of the questionnaire were developed. Each version used one of the following branching instruction designs.

Experimental Branching Instruction Designs

The Control Method. Shown as the first design on the left in Figure 1, this is the branching instruction used by the U.S. Bureau of the Census, with the check boxes on the left and the response options on the right. The branching instruction is placed to the right of the response option with no change in size or font from the rest of the text (which is 10-point Frutiger), except that the instruction is in Italics rather than normal print.

The Detection Method. In this method, which is illustrated by the middle design in Figure 1, the check boxes and the branching instruction are in the same location as on the control. However, the branching instruction is enlarged and boldened to attract respondents' attention to it. Also, a bold arrow comes off the non-branching 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-branching choices, e.g., "(If yes)."

It is hypothesized that if a respondent chooses a response associated with a branching instruction, they will be more likely to perceive the branching instruction in the detection method than the control, whereas if they choose a response devoid of a branching instruction, their eyes will be drawn to the next question more often in the detection format than the control because of the left-hand arrow.

The Prevention Method. The method shown on the right of Figure 1 includes an instruction to pay attention to the branching instructions. Furthermore, the position of the check boxes and response categories are reversed, which makes it possible to place the branching instruction immediately beside the check box and presumably within the view of respondents. Second, the branching instruction is enlarged. Third, the background is changed from yellow to white to increase the contrast between the bold lettering and the background even further.

It is hypothesized that if a respondent chooses a response with a branching instruction associated with it, they will be more likely to perceive the branching instruction in the prevention method than the control, whereas if they choose a response devoid of a branching instruction, their eyes will be as likely to go to the next question in the prevention format than the control.

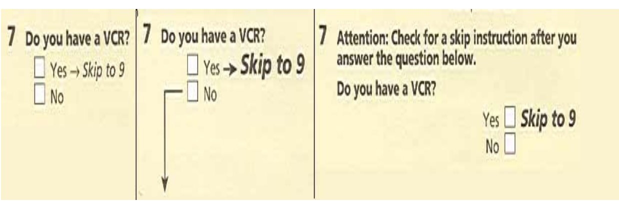

Figure 1. Branching instruction designs.

<u>Respondents</u>

Respondents were recruited in the fall of 1998 through the use of fliers or by word of mouth and paid $10 to participate in the study. Eight respondents were randomly assigned to the first treatment, nine to the second, and eight to the third, for a total of 25 participants. The number of males (12) completing the study was nearly equal to the number of females (13). The mean age of respondents was 40. Since one of the goals of this research is to improve the design of questionnaires for the less educated, half of the respondents had less than a high school education. Also, a concerted effort was made to include a mixture of races and ethnicities (Hispanic, Asian, black, and white) in the study.

<u>Collecting Eye Movement Data</u>

The photoreceptors of the human retina show a pronounced density peak in a small region known as the fovea. In this region, which subtends a visual angle of about 1°, the receptor density increases to about 10 times the average density. The nervous system controls the muscles attached to the eye to keep the image of current interest centered accurately on the fovea because this results in the highest resolution image. The appearance of high resolution at all directions outside of this region is an illusion maintained by a combination of physiological mechanisms (rapid scanning with brief fixations), and psychological ones. For example, there is a blind spot on the eye where no photoreceptors exist; however, the brain compensates for this. Also, a character on a typical computer display screen or piece of paper subtends an angle of about 0.3°, or roughly 3 millimeters at a normal viewing distance of 60 centimeters. Such characters cannot be accurately resolved unless the eye is aligned for 0.05 seconds.

As shown in Figure 2, the fovea is in line with the optical axis. The cornea is the curved portion of the eye in front of the lens. The pupil is the opening of the eye. Light passes through the cornea, pupil, and lens of the eye, focusing on the retina, specifically the fovea. The iris, seen as the colored part of the eye, controls the pupil size.

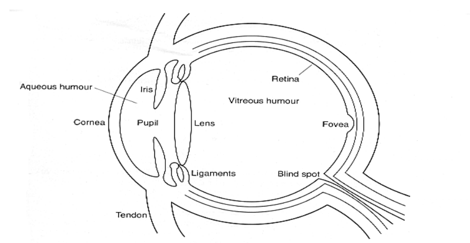

Figure 2. The human eye.

In this study, the **E**ye-gaze **R**esponse **I**nterface **C**omputer **A**id (ERICA) system was employed to collect the eye movements (Hutchison et al. 1989). An infrared light emitting diode (LED) resides at the center of a lens that is attached to a camera feeding its signal to a computer. This LED bathes the user's face in infrared light, a wavelength of light invisible to humans. When the light from this properly positioned LED strikes the eye, two features become apparent to the camera, as shown in Figure 3

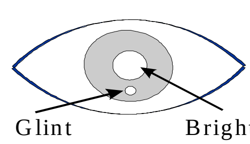

Figure 3. The two features formed when infrared light strikes the eye.

The glint is the specular reflection of the LED itself off the cornea. Essentially, a specular reflection is the intense reflection of light off a curved, shiny surface.

The glint appears as a small, bright dot to the camera. The bright eye is the absorption and reemission of the infrared light by the retina of the eye. To the camera, this makes the pupil glow

The ERICA system locates these two features of the camera image and determines where the user is looking based upon the separation between the two features, as shown in Figure 4.

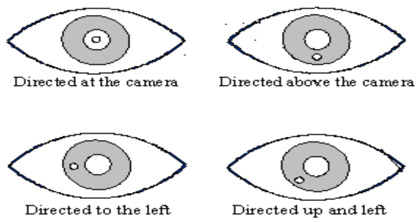

Figure 4. Glint and bright eye relationships. *The separation between the glint and the bright eye allows computation of gaze position.*

Figure 5 shows the hardware used for this study. A cardboard box with a clipboard mounted to it rested where the computer monitor normally resides. The questionnaire was mounted vertically on the clipboard. The eye-tracking hardware, which included the camera, LED and mirrors, was housed in a module that rested beneath the clipboard arrangement. The mirrors directed the camera's line of sight to the respondent's eye. To the left of this set-up, a 9.5 inch black and white monitor displayed the respondent's eye and glint. Behind the respondent, a computer monitor displayed the respondent's eye-movements while they were in the process of being recorded (not shown in Figure 5).

It was discovered during a pretest of this set-up that eye tracking data was not recorded if respondents sat too close to the questionnaire because their eyes had to undergo too large a degree of rotation to look at the top of the questionnaire. Positioning respondents from 40-45 cm away from the questionnaire lessened the degree of eye rotation to an acceptable level while still allowing the respondent to sit close enough to the questionnaire to mark their answers.

Respondents with contact lenses and eyeglasses participated in the study; however, respondents with eyeglasses had to be positioned such that extra reflections off their eyeglasses caused by the infrared LED did not obscure the pupil of the eye. These respondents needed to lean their head back on the chair, as shown in Figure 5, for their eye movements to be successfully captured.

Figure 5. Equipment Setup.

<u>Experimental Procedure</u>

The respondent filled out a screener questionnaire that established his or her demographic profile. Then the respondent was seated, and the experimenter situated him so that the system could observe his eye and the respondent could reach the questionnaire comfortably. Next, the respondent was calibrated while reading a test page of printed text. If the respondent's eyes were not tracking accurately, his eye movements were recalibrated. All respondents successfully completed calibration before starting the questionnaire. Typically, inaccuracies resulted from the respondent moving his head too much in the beginning; thus, the test page served to acclimate the respondent to the system setup and the constraints on his motion – respondents could only move their heads about 5 centimeters in any direction and still have their eye-gaze data captured.

When the eye-tracking data looked good, the respondent began to fill out the questionnaire. The experimenter needed to reposition respondents if their eyes showed signs of not being captured anytime during this process.

<u>Analytic Technique</u>

A question's structure was parsed into the following four components: (1) the question, (2) the answer categories, (3) the check box, and (4) the branching instruction. In addition, return sweeps between question components were examined. A return sweep occurs when the eye moves from the end of one line of text to the beginning of the next. Since the accuracy of the system is roughly 1 centimeter and the text of the questionnaires is dense, it is impossible to reliably identify which question component the respondent is observing by examining a single gazepoint; however, by examining a series of gazepoints or a cluster of gazepoints in relation to each other and the printed information on the questionnaire, the gazepoints were feasibly matched to the question components, as shown in Figure 6.

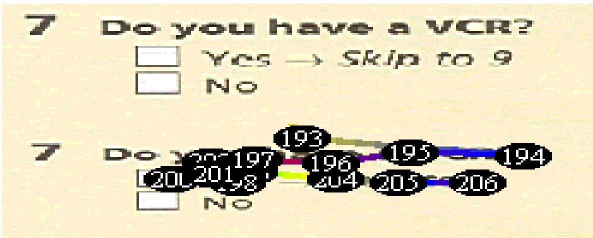

Figure 6. Gaze trail superimposed on a question and then parsed. *Gaze trails 193-194 corresponds to observing the question. Gaze trail 194-197 reveals a return sweep of the eye. Gaze trail 197-203 correspond to observing and marking the answer category, and gaze trail 204-206 corresponds to looking at the branching instruction.*

An *error of commission* occurs when respondents are instructed to branch ahead to a specified question, but instead they answer the next question, or some question in between. An *error of omission* occurs when respondents are supposed to answer the next question, but instead branch over it.

Results

All respondents completed calibration and answered one version of the questionnaire. The eye tracker never failed to reacquire a respondent after losing track of their eye; however, every respondent had a gap in their data where tracking failed. This usually occurred because respondents leaned forward in their chair to mark their answers and moved out of the field of view of the eye-tracking camera. Also, tracking sometimes stopped because respondents would raise their head or the rotation of their eyes was too great when they looked at the top of the questionnaire. However, only one respondent exhibited data that could not be analyzed in any way due to too many data gaps. This was most likely due to a mis-calibration that the experimenter did not correct. Overall, about 45% of the analyzed questions had answers with no data gaps, while 85% to 90% of the eye movement data was interpretable, a promising result considering the exploratory nature of the study.

Executing a Branching Instruction Correctly

Figure 6 shows a typical respondent's perception order for a question with a successful execution of a branching instruction. Generally, the respondent would read the question, observe the answers, fill in the answer, and observe the branching instruction. Alternatively, the last two steps were reversed: the respondent observed the branching instruction and filled in the answer, as shown in Figure 7. These eye movement patterns suggest that respondents are more likely to execute the instruction correctly if they read it very near to the moment when they move to the next question.

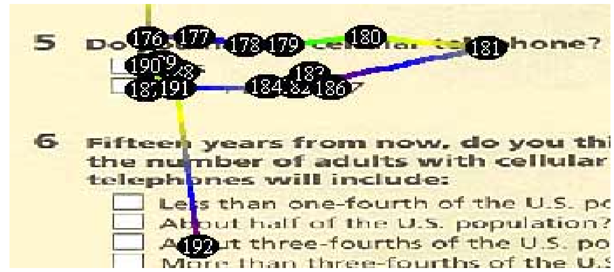

Figure 7. Alternative gaze trail of a respondent executing a branching instruction correctly. *Perception order is question and return sweep (176-182), the branching instruction (183-186), and response categories (187-191). The respondent marked 'no' in response to question 5*

Errors of Commission

Errors of commission occur when respondents are supposed to branch to a specified question, but do not. A corollary trend emerged from the eye movement patterns of respondents making errors of commissions from those who executed the instruction correctly; that is, respondents tend to make errors of commission if they do not observe the branching instruction immediately prior to or after marking their answer. Either they never perceive the instruction at all, as illustrated in Figure 8.

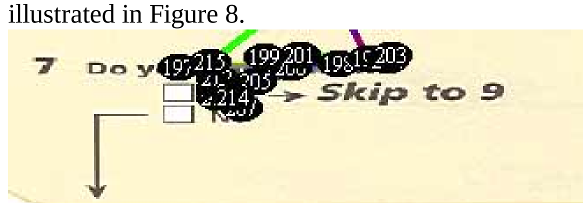

Figure 8. Failure to observe the branching instruction. *Perception order is question (197-204 and answer (205-215). The respondent marked 'yes' in response.*

Or, they see the branching instruction prematurely, as illustrated in Figure 9.

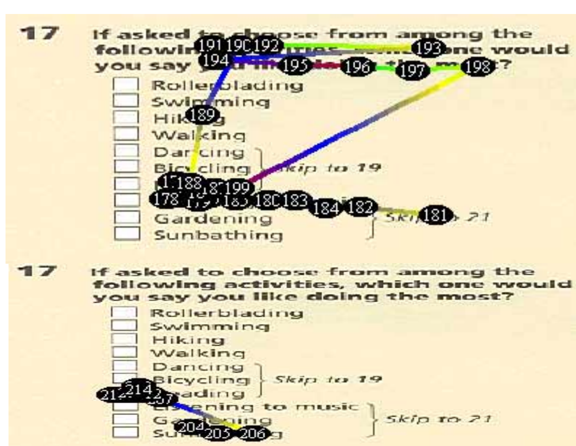

Figure 9. Premature observation of a branching instruction. *At the end of reading this question, the respondent observed an answer (178-180), the branching instruction (181-182), the answer again (183-188), the question (189-198), other answers (199-206), and the final answer (207-214). The respondent marked 'listening to music.'*

The eye movements shown in Figure 9 suggest that if respondents read the branching instruction too early in the process, they fail to recall it later. Consequently, they go on to the next question in the series rather than branching to the one specified in the instruction.

Errors of Omission

Errors of omission occur when respondents are supposed to answer the next question, but do not. The eye movement patterns suggest that these errors occurred for two different reasons. In the one instance, respondents looked at the instruction when it did not apply to them and erroneously executed it. In these instances, the respondents typically observed the branching instruction last, as shown in figure 10. This result reinforces the previous findings--that reading the branching instruction during the critical timeframe when respondents are preparing to move to the next question (i.e., last) determines whether they execute it. It also suggests that reading the instruction during this time frame is good when the instruction is associated with their answer choice and bad when it is not.

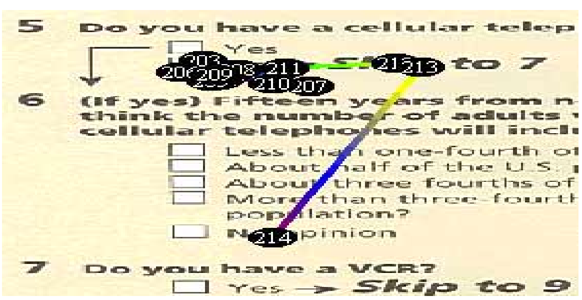

Figure 10. Typical eye gaze trail of a respondent making an error of omission. *The respondent observed the branching instruction last (gazepoints 212-213) and executed it, despite marking the Yes answer.*

However, the eye movements reveal another reason respondents did not answer the next question when they were supposed to. An example of this is given in Figure 11 in which a respondent marked the response category 'country,' which required no branching. After reading the remainder of the response options, the respondent advanced to question 22, as he should have, and read the question there. However, rather than answering question 22, he proceeded on to 23, possibly because question 22 required generating an answer, which is a more difficult task, rather than selecting an answer from among a preprinted list. Without the aid of the eye-movement data, it might appear as though the respondent executed the branching instruction in question 22 (to advance to question 23). However, the eye movements reveal that the respondent never read that instruction

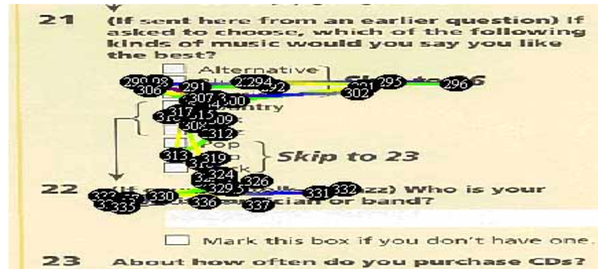

Figure 11. Typical gaze trail of what appears to be item non-response

Discussion

Despite the fledgling nature of the eye movements collected in this study, two findings are evident. One is that respondents make errors because they do not perceive the branching instruction. The other is that the order in which respondents read information affects whether or not they understand it as intended. Essentially, there appears to be a critical moment in the process of navigating from one question to the next—either immediately before or after marking a response—that respondents are receptive to acting upon the branching instruction. If they read it earlier than this, they fail to recall it when it comes time to use it. Alternatively, during this timeframe, they may mistakenly act upon it as soon as they read it.

This finding implies that a kind of visual grammar or syntax occurs that affects respondents' understanding of the task or performance, and it provides direct empirical evidence for a hypothesis which previously had only indirect evidence or face validity in its favor: that is, answering a visually administered questionnaire is fundamentally different than answering an interviewer-administered questionnaire because respondents are free to select how to view the information presented to them (Jenkins and Dillman 1997). Thus, this finding lends credence to the assertion that we need to understand the perceptual process well enough to exert control over it.

At the beginning of this study, it was hypothesized that different branching instructions designs would affect respondents' eye-movements, and although there is evidence for this from the error rates across the three designs in this study, the eyetracker was not sensitive enough at the time of this study to enable this conclusion to be derived from the eye-movement data itself. However, since this study was performed, notable advances have occurred in the system, which may facilitate this analysis in the future. For instance, the system can now identify where someone is looking 60 times a second as opposed to the maximum of 15 times a second when this study was performed.

In addition, system limitations necessitated that respondents read and answer a questionnaire that was placed vertically (i.e., at a 90-degree angle), which is not representative of natural form-filling behavior. A follow-on study is now underway in which it has been possible to mount the questionnaire at a 45-degree angle.

Changes are also being made to the software to facilitate automatic coding of a subject's eye-movements. Correlating these patterns to respondent behaviors can lead to a scientifically rigorous protocol for judging the behaviors independent of any potential prejudices of the researcher.

Finally, plans are in place for expanding the application from paper to the Web. This equipment was originally designed for use with a computer monitor, so it should be easier to use in a Web environment than paper. However, the quantity of eye-movement data produced on the Web will be just as prodigious as for paper, and consequently learning how to automatically code the data will be as important for the Web as it is for paper.

Conclusions

Eye-movement analysis does appear to be a promising new tool for evaluating visually administered questionnaires. Hardware and software that was originally developed for use with computer monitors was adapted to track the eye movements of respondents answering three versions of a paper questionnaire, which differed in the visual designs of their branching instructions. The study revealed that respondents were more likely to execute the instruction correctly if they observed the instruction immediately prior to or after marking their answer compared to reading the branching instruction prematurely or not reading it at all. This is an insightful finding and an encouraging lead, one that could not be drawn from any other method.

Although the eyetracking equipment and methods have not been perfected, there is every reason to believe they will improve with time and, more importantly, that the benefit derived from this methodology will extend beyond respondents' understanding of branching instructions to other information on the questionnaire--for instance, the questions themselves. Work should continue towards improving the methodology and applying it to other areas of interest in the future.

Acknowledgements

This paper reports the results of research and analysis undertaken by Census Bureau and the University of Virginia staff. It has undergone a more limited review than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

References

Ellis, S., Candrea, R., Misner, J., Craig, C.S., Lankford, C.P., Hutchison, T.E. 1998. Windows to the Soul? What Eye Movements Tell Us About Software Usability. Paper presented at the Usability Professional's Association, Washington DC.

Featherston, F. and Moy, L. 1990. Item Nonresponse in Mail Surveys. Paper Presented at the International Conference on Measurement Errors in Surveys, Tuscon, Arizona.

Hutchinson, T. , White, K. , Martin, W., Reichert, K., and Frey, L. 1989. Human Computer Interaction Using Eye-Gaze Input, *IEEE Transactions on Systems, Man, and Cybernetics* 19(6):1527-1534.

Jenkins, C. and Dillman, D.. 1997. Towards a Theory of Self-Administered Questionnaire Design. In L. Lyberg et al. (eds.) *Survey Measurement and Process Quality*. New York: Wiley Interscience. Pp. 165-196.

Lohse, G.L. 1997. Consumer Eye Movement Patterns on Yellow Page Advertising. *Journal of Advertising*. 26, (1): 61-73.

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

Rayner, K. 1992. *Eye Movements and Visual Cognition: Scene Perception and Reading.* New York: Springer-Verlag.

Rayner, K. 1983. *Eye Movements in Reading: Perceptual and Language Processes*. Academic Press, Inc.

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

Redline, C., Dillman, D., Smiley, R., Lee, M. and DeMaio, T. 1998. Beyond Concurrent Interviews: An Evaluation of Cognitive Interviewing Techniques for Self-Administered Questionnaires. *Proceedings of the Section on Survey Research Methods,* American Statistical Association, pp. 900-905.

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 et al. (eds) *Survey Measurement of Drug Use Methodological Studies,*. Washington, DC. National Institute of Drug Abuse, U.S. Department of Health and Human Services, 221-243.