ISER Working Paper Series

Real-world Eye-tracking in Face-to-face, Web and SAQ Modes

Olena Kaminska
Institute for Social and Economic Research
University of Essex

Tom Foulsham
Department of Psychology
University of Essex

No. 2013-07
June 2013

INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

www.iser.essex.ac.uk

Non-technical summary

The same question administered to the same group of people may receive different estimates depending on whether an interviewer administers an interview, or a respondent answers it via web or fills a paper and pencil questionnaire. This dependency of answers on survey modes have been recognized for some time now, yet little is understood what causes such differences. The current research provides a glimpse into respondents' brains when they answer survey questions. Specifically, we explore what new insights into understanding survey modes differences can be obtained via an innovative method of tracing respondents' eye-movements. We measure respondents' visual attention to question wording and response options and compare these across three modes of data collection: face-to-face interviewing with response options sheets (show cards), paper and pencil self-completion and web self-completion.

To record respondents' eye-movements we asked 20 volunteers to wear a head device which records the visual view of a respondent, and aligns with it respondents' eye fixations. Thus, we can know with some precision where and how long respondent was looking at any point during an interview. A coder then watched the recordings and coded the beginning and end of each glance at different attributes of questions, such as question wording and each response option. We were able to trace and code eye-movements in each of the three modes successfully.

Such new information provides many details on how respondents read survey questions and how this compares across modes. We found that in web and paper and pencil modes respondents often reread question wording. In face-to-face mode when an interviewer reads question wording respondents simultaneously read response options from show cards at least part of the time. When a response scale is presented vertically respondents do not always read the bottom response options in all modes, but reading the bottom options is more common in face-to-face mode. Interestingly, not all response options are read in order as presented, from top to bottom. But reading response options in order is more common in web and paper and pencil modes than in face-to-face: as many as 8% of questions in a face-to-face mode had readings of response options from bottom to top.

Overall, we find that studying eye-movements to understand the question response process may be very useful, especially in finding causes of survey mode differences.

Real-world Eye-tracking in Face-to-face, Web and SAQ Modes

Olena Kaminska+
Institute for Social and Economic Research, University of Essex
Tom Foulsham++
Psychology Department, University of Essex

+ Corresponding author: Institute for Social and Economic Research, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, UK (email: olena@essex.ac.uk)
++ Psychology Department, University of Essex, UK (email: foulsham@essex.ac.uk)

Abstract: Eye-tracking is becoming a popular testing tool to understand how different forms of asking questions influence respondents' answers. Until now, due to the ease of eye-tracking on PC, this method has almost exclusively been used to test questions in web/PC mode. Our paper extends the application of eye-tracking to other modes with a visual component, including self-administered paper and pencil (SAQ) and face-to-face with show-cards (PAPI) in a realistic interview set up. The current article reports on a feasibility study of using a real-world eye-tracker to measure visual attention in the modes with visual component.

Key words: eye-tracking, mode effect, primacy effect, show cards.

JEL codes: C81, C83

Acknowledgements: This study was funded by the University of Essex Interdisciplinary Research Fund for Social Sciences "Understanding Survey Responses through Eye-Tracking: Mixed-Mode Comparison" granted in 2011 (Principal Investigator: Olena Kaminska; co-PI: Tom Foulsham). The authors are grateful to Gemma Luxton for hours of diligent coding of eye-tracking data.

Introduction

In general, the survey response process consists of four main stages: comprehension of the question and question instructions, retrieval of relevant information, judgment, and reporting of the survey answer (Tourangeau, Rips, and Rasinski 2000). Nevertheless, the same questions will not always go through the same response process under different conditions. Survey modes have been known to influence survey answers even if the wording of the question, question instruction and response scales are the same (Bowling, 2005; de Leeuw, 2005).

While mixed mode studies can have gains in response rate, survey coverage and cost (Groves et al., 2004; de Leeuw, 2005), there is a resistance to using mixed mode studies because of mode effects. Their presence means that different sample members are measured with different quality. Many studies have explored the differences between modes (for overview see Bowling, 2005; de Leeuw, 2005) and some questionnaire design features which decrease these differences (Dillman, Smyth, and Christian 2009). Nevertheless, some differences and their actual causes remain unexplained which opens an avenue for innovative survey response testing tools.

A number of methods are used to study the survey response process with the aim to understand better survey error cause and to improve questionnaires. Cognitive interviewing, when a respondent reports himself how he arrived at an answer, is a common questionnaire pretesting tool. Such techniques as concurrent and retrospective verbal probing and think aloud are used to understand the nature of measurement error (Willis, 2005). Issues with misinterpreting question instructions and question wording, with recall, with response mapping on response scale and visual layout are often revealed in cognitive interviewing, and as a result a questionnaire is improved before it is fielded. Audio or video recording of survey interviews and subsequent coding of interviewer-respondent communication has also been used successfully in studying interviewer effects on the response process (Jans, 2009) or comparing response process in different ways of question administration (Bilgen and Belli, 2010). An unobtrusive computer audio-recording (CARI) has also been used to detect field-interview fabrication and has been recently shown to be useful for quickly identifying potential for measurement error (Hicks et al., 2010).

Eye-tracking, a method which records eye movements and aligns them with the visual picture that respondent is looking at, is an innovative approach to studying the survey response process. It was first introduced by Redline and Lankford (2001) and Graesser et al (2006), and since Galesic et al.'s (2008) publication, which studied the nature of primacy effects in web mode, has gained popularity as a tool to explore visual attention (e.g. Ashenfelter and Hughes, 2011; Kamoen et al., 2011; Lenzner et al., 2011; Menold et al., 2011; Romano, and Chen 2011). Eye-tracking has now received recognition as a successful pre-testing tool, with its current regular use by United States Census Bureau (Ashenfelter and Hughes, 2011; Romano, and Chen 2011) and the German Federal Statistical Office (Federal Statistical Office, 2011).

To date, due to its technical ease of implementation on PC (Galesic and Yan, 2011), eye-tracking has been almost exclusively used in web surveys. Two notable exceptions include feasibility studies of eye-tracking in SAQ mode (Redline and Lankford, 2001; Potaka, 2007). While the studies demonstrated advantages of tracing eye-movements for SAQ questionnaire development, they lacked a realistic set-up for the interview environment: the paper questionnaire was placed vertically on a clipboard, and respondents had to be seated at an unnatural distance from the questionnaire. Our study for the first time extends the use of eye-tracking to other modes with a visual component, including face-to-face with show-cards and SAQ mode in a natural context.

We report on a feasibility study of using eye-tracking methodology for untangling differences in response process across survey modes. Recognizing that survey context has a potential of influencing respondents' answers (Schwartz, 1996), we pay attention to keeping survey response environment as close to the real survey situation as possible; and we demonstrate that the new real-world eye-tracking technology enables tracing eye-gazes without restrictions to respondents' posture or position of a survey instrument.

The first section provides an overview of the current eye-tracking usage in survey methods research and an introduction to real-world eye-tracking. After a short description of data collection for the feasibility study, we report on success and challenges of using real-world eye-tracker in each specific mode, and provide recommendations for future studies using this technique. The following section provides an example of how eye-tracking information can be used to compare response process across modes. Specifically, we explore differences in time spent on question wording and response options and the order of their viewings across modes. Although the section's main purpose is to demonstrate detail level and topics that are possible to study via real-world eye-tracking, its findings are also interesting from methodological perspective and are suggestive of the future research avenues. Finally, the discussion section summarizes the findings.

Eye-tracking in Survey Research

In designing questions for an interview with visual component we often expect that a respondent pays close attention to each word of a question and each scale point. Such expectations are nevertheless naïve and they do not seem to represent the reality of how we read text or how we perceive question and scale points' layout.

Previous research in the field of Psychology has confirmed that some visual items attract our attention more than others. For example, bright or unusual objects gain longer viewer fixations (Underwood & Foulsham, 2006). In the context of text, we read progressively, skipping highly predictable words, fixating longer on unusual or low-frequency words, and rereading ambiguous or unfamiliar words (Rayner, 2009). Our visual attention may vary, with more words being skipped when subjects admit paying little attention to the task (Reiche et al., 2010). We also pay little or no attention to advertisement banners on web pages, and eye-tracking has been used to study how changing location and look of web page banners can attract visual attention of viewers (e.g. Burke et al., 2005).

A consistent pattern of eye-movements has been observed when respondents read survey questions in a web mode. Respondents' eye fixations are prolonged by such linguistic characteristics of survey questions as low-frequency words, vague or imprecise relative terms, vague or ambiguous noun phrases, complex syntax, complex logical structures, and low syntactic redundancy features (Lenzner et al., 2011a). Negative wording of a question or response options receives higher chance of rereading it in comparison to positive wording or the same question or response option (Kamoen et al., 2011). Just as web banners are not attended to very often, in survey research respondents are found to pay little attention to progress indicators, and help links (Ashenfelter and Hughes, 2011).

In a small scale study at the University of Maryland Galesic et al. (2008) found that around 10% of respondents never looked at the last two response options in a 12-category list of children characteristics, and at least one response option was skipped by around 75-80% of participants across different questions. In addition, the authors showed that spending more time looking at response options in the first half of a list results in selecting a response among these options (primacy effect). Understanding that visual attention varies when reading survey questions can help us to design them better by avoiding word features which increase cognitive effort and by bringing attention to those question features and instructions which are most important.

Eye-tracking is especially useful in exploring how changes in the visual or syntactic part of a questionnaire may improve respondent's attention when reading survey questions. A few recent studies have compared eye-movements and timing of eye fixations when presented with different layout of the same questions. By comparing fully-labelled and end-labelled response scales, Menold et al. (2011) found that there is no significant difference in fixation time between the scales overall, but respondents spend longer time looking at the most right handside label in end-labelled scales than in fully-labelled scales where respondents tend to spend time looking at other categories as well. Another experimental study compared eye-movements when response boxes are placed to the right versus to the left of the response scale options (Lenzner et al., 2011b). The authors discovered that answer boxes to the left of the response options decrease response latencies, decrease number of fixations on answer boxes and decrease number of gaze switches between answer boxes and response options.

The only two studies which used eye-tracking for SAQ questionnaire both explored eye-movements when respondents read Census forms. The study in the US by Redline and Lankford (2001) discovered that respondents notice routing instructions depending on their location with respect to previous question. This article encouraged further experiments which led to an improvement in routing instructions (Redline and Dillman, 2002). A study from New Zealand by Potaka (2007) found that respondents notice only the first bubble instruction indicating how to mark answers, and skip the later ones. In addition the author showed that respondents do not look at some of the response options which appeared below alpha-numeric box, but are more likely to notice them if they actively search for such response option.

Real-World Eye Tracking

Recent eye-tracking studies are largely limited to exploring response process in web mode. This is mainly due to technical ease of tracking eye-movements when respondents look at a PC screen. The lack of eye-tracking studies in SAQ may be explained by an unnatural position of the paper questionnaire in previous studies (Redline and Lankford, 2001; Potaka, 2007). But just as the authors of the latter studies have hoped for, the technology has developed since to allow tracking eye-movements in SAQ in a more natural set up. Our study for the first time extends the use of eye-tracking to face-to-face mode with show cards and to SAQ where questionnaire is positioned naturally.

We therefore used an innovative set up with a real-world eye tracker (the SMI HED system). This equipment monitors eye movements using a camera worn on a headset, while a second camera captures the scene in front of the participant as well as any sounds. The system is calibrated by asking the respondent to look at a sequence of points on a wall or desk so that eye fixations are mapped onto points in the real world. Participants are free to move around, their view is not limited to a computer screen, and eye gaze can be recorded on anything in their field of vision. Because this field of view changes with each head movement, and is not constant like the computer monitors used in previous studies, additional coding is required if a researcher wants to identify what is being looked at in different situations.

Real-world, also called mobile, eye-tracking is a relatively new tool that has been used successfully to study vision and attention during such diverse behaviours as walking, driving and playing cricket (Hayhoe & Ballard, 2005; Land & Lee, 1994; Land & McLeod, 2000). Studying visual attention in these contexts has revealed the coordination and timing involved in real world actions—such as fixating an object before reaching to pick it up. Moreover, while much can be learned with experiments on a computer screen, visual attention in realistic environments may be different. For example, Foulsham, Walker and Kingstone (2011) found that people looked at different locations when watching a video of a street as opposed to walking down the same street in the real world. Freeth, Foulsham and Kingstone (2013), meanwhile, found that attention to a video of an interviewer asking questions was different from when the same questions were asked face-to-face. These differences make it clear that it is important to apply eye tracking methods to a range of realistic survey modes.

Method

We report on a feasibility study which was designed specifically to explore whether real-world eye tracking can be used to study the survey response process in each of the three modes: web, face-to-face with show cards and SAQ; and whether comparable fixation data can be obtained for comparison of visual attention across the three modes. It was decided to use real-world eye-tracker for web mode as well in order to obtain comparable measures.

The feasibility study was carried out in two stages as part of collaboration between the survey methods team and the Psychology department in a UK University. First, we explored the best set up for each mode separately on 5 participants. A number of meetings between experts in Psychology and Survey Methods took place to find the closest set up to real survey interview while obtaining good quality eye tracking information. At the second stage an additional 20 participants took part in a small experiment where each respondent was randomly assigned to one of the three modes. Of 14 respondents, whose data was used in analysis, 21% were male, 21% were not native English speakers, and all were between 20 and 26 years old.

All participants were interviewed in a Psychology Lab using a real-world eye tracker (the SMI HED system; Sensorimotoric Instruments; Teltow, Germany). The videos of the visual view of respondents with aligned eye-fixations and sound (e.g. interviewer and respondent speaking) were recorded. Respondents were not restricted in their head movements or instrument position when wearing the real-world eye tracker, and were not given any instructions on posture or placement of survey instruments. Participants were selected from volunteers for Psychology experiments, with the requirement of having normal vision. Each interview took around 15 minutes, with approximately 2-5 minutes of eye-tracker set-up. All respondents were paid £6 for participation.

The survey questions selected for this study were previously administered to panel members of a large national survey panel in two modes, in SAQ and CASI (equivalent to web). The question font and layout was not changed in these modes, and show cards (of A4 size) for face-to-face questionnaire were developed based on these questions. Additional demographic and personality questions were asked separately in web mode in the end of the experiment.

One section of the questionnaire, comprising of 12 general health questionnaire items (GHQ-12, see Goldberg and Williams, 1991), was selected for manual coding. All questions have 4 point fully-labelled scale presented vertically and they were placed in the middle of the interview. Using custom annotation ChronoViz software (Fouse et al., 2011) one coder noted the beginning and end time of gazes at question wording, at each response category, and where this was not possible, at the top two or bottom two categories. Respondents also looked away from the survey materials (including at the interviewer in face-to-face mode, or other objects in the room) which was coded as 'looking at other'. Finally, 13-14% of gazes across modes were uncodable. From audio and video recordings the start and end times of the interviewer reading the questions were also recorded. Due to limited quality of 6 recordings from SAQ mode, only 2 SAQ interviews were coded, leaving a total of 14 interviews (additionally 7 in web and 5 in face-to-face mode) suitable for analysis. The number of completed interviews per mode is still sufficient for statistical analysis, as demonstrated in the next section, given that analysis is at the question level with control for clustering within respondents.

Using real world eye-tracking for mode comparison

We were successful in tracking eye-movements while administering questionnaire in each of the three modes (SAQ, web and face-to-face with show cards). Figure 1 presents example video shots from our study on how real-world eye-tracker tracks eye-fixations in each mode.

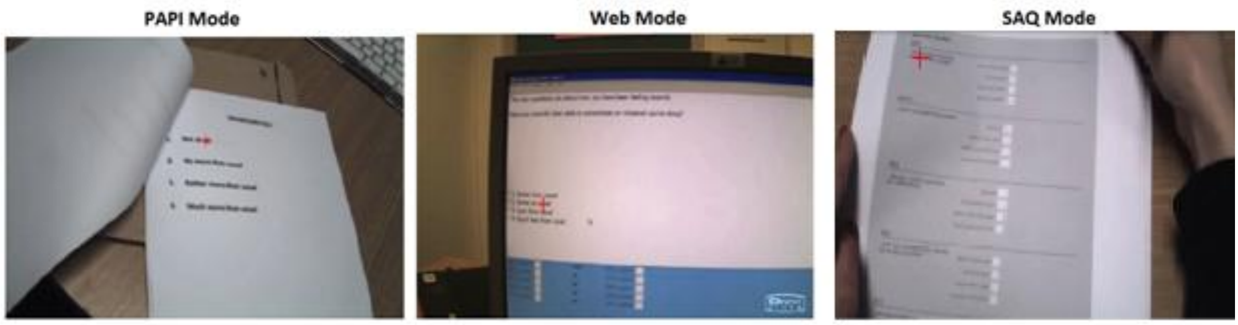

Figure 1. Examples of eye-tracking in face-to-face, web and SAQ modes

Web was probably the easiest and most consistent mode for tracking eye-movements using real world eye tracker. This is due to the limited area of view, and as a result fewer head movements. An important limitation of the real-world eye-tracker in comparison to the eye-trackers used previously in web mode is the requirement for manual subsequent coding of time span for each eye fixation. Nevertheless, this tool has an advantage of tracing gazes away from the PC screen such as at the keyboard, mouse, mobile phone, or any other distractions.

Tracking eye-movements in a face-to-face mode with show cards appeared to be the second most successful. Our original aim was to track eye-movements on both the show cards and an interviewer such that we could differentiate when respondents look at an interviewer and when they look at show card including which category they look at. Because the eye-tracker has a smaller field-of-view than do humans, and because respondents usually switched gazes between the interviewer and a show card with an eye-movement without moving their heads, it was often not possible to track both simultaneously. Nevertheless, we could successfully set up eye-tracker for show-cards or an interviewer separately. Because we are mainly interested in eye-movements while respondents read show cards, for the second part of our study we tracked eye-movements on show cards only. We know when respondents were looking away from the show card, and we can usually infer when it is a look at an interviewer (e.g. when a respondent asks for clarification or when an interviewer reads a question). In the future studies with an interest in gazes at an interviewer the eye-tracker can be set up to view an interviewer instead.

SAQ was the most challenging of the three modes. Our aim was to trace eye-gazes in the realistic to the survey SAQ interview situation, i.e. allowing for full flexibility of respondent to hold / place / and fill questionnaire in any way that they like. We also wanted to keep the situation private by leaving the respondent by themselves in the room. This prevented us from interfering in order to adjust the eye-tracker. Providing full flexibility during SAQ interview appeared to be particularly challenging with the first 6 interviews in SAQ mode resulting into only partially codable data. The common problem was that after a few first questions the quality of tracking decreased sharply due to changes in the position of the questionnaire or equipment. After careful review of videos we realized that as soon as an interviewer left the room, respondents changed the posture drastically, usually by slouching. This meant that the eye-tracker was set up for a different angle than the angle of questionnaire completion. A surprisingly simple solution was found. We asked respondents to start filling the questionnaire first allowing respondents to take the natural posture. After the first 2 questions were filled the eye-tracker was set up, and the interviewer would leave the room. Two interviews were conducted using this approach, both of which resulted in similar quality to other modes. Importantly, we succeeded in tracking eye-gazes without setting any restrictions to respondents and by ensuring privacy as well.

Overall, we find that in each of the three modes it is possible to measure the visual attention of respondents well enough to differentiate whether they are looking at an instruction, at the question wording, or at the top or bottom part of the response scale. For many questions it also is possible to differentiate exactly which response option the respondents are looking at. This is achieved with no restrictions or instructions to respondents on where they should look or how to move their head. We found that for almost all situations when respondents move show cards or SAQ questionnaire in different positions we are able to track their eye-movements. The only restriction is that participants had to wear the tracking helmet with a small camera and system on it. Importantly, the system has been used in real world situations (e.g. while walking around campus and ordering coffee) and has the potential to be used outside of the Psychology Lab as part of a real survey interview. Additionally, the latest systems are considerably less restrictive and can fit onto a pair of glasses.

Empirical example: Comparing visual attention between modes

To-date there has been no study which compares visual attention of respondents across face-to-face, SAQ and web modes. We therefore know little how respondents perceive survey questions in SAQ, do not know whether and when respondents read show cards in face-to-face interview, and how the attention to survey questions and response categories differs across modes. This section provides the first empirical insights into cognitive response process differences across modes using eye-tracking. We compare modes using multiple or logistic regression at question level with correction for clustering within respondents. All models are run in Stata 12.1.

1 Attention to question wording

Overall, respondents take significantly longer time to listen to each question and read response options from show cards in face-to-face mode than to read question and response options in self-administered modes (Table 1, Model1). We find that on average the total time spent per question in face-to-face mode is 8.7 seconds, which is 2.6 seconds longer than in web, and 3.5 seconds longer than in SAQ. One may suggest that such result is expected given that the time for reading a question out loud should take longer than reading a question in one's mind. But this does not seem to be true if we look at the total time spent on the question wording (in self-completion modes this includes rereadings). The total time spent on a question in face-to-face mode is not significantly different than in web mode (Table 1, Model 2). Even though the time spent on question is shorter in SAQ mode than in face-to-face mode, this difference is only 1.2 seconds on average, and there is still 2.3 seconds of overall time difference per question that remains to be explained.

Table 1. Mode effect on total time per question, question reading / listening time for all questions and for questions read only once

	Model 1, Total time			Model 2, Total question time			Model 3, Total question time if read only once		
	Coef.	Linear. SE	p-val	Coef.	Linear. SE	p-val	Coef.	Linear. SE	p-val
Intercept	8780	506	<0.001	2872	139	<0.001	2872	139	<0.001
web (PAPI)	-2560	982	0.02	-76	191	0.70	-957	229	<0.001
SAQ (PAPI)	-3484	1492	0.04	-1219	190	<0.001	-1627	184	<0.001
N questions	168			157			96		
N respond's	14			14			14		
Model Fit	F(2,12)=4.82, p=0.03			F(2,12)=24.99, p<0.001			F(2,12)=35.94, p<0.001		

*Total time indicates time spent on the question overall (listening to the question and reading response options in face-to-face mode; reading question and response options in SAQ and web modes). Time is presented in milliseconds

Does this mean that respondents read questions in self-completion with a similar speed as an interviewer reads them out loud? Interestingly, the effect appears to be due to a number of rereadings in self-completion modes. In face-to-face mode interviewers read each question only once. In contrast, in self-completion mode around 2/5 of questions (37% in web and 39% in SAQ) are read only once, another 2/5 of questions (39% in web, and 43% in SAQ) is read twice, and the rest is read three times or more. In fact, in a web mode the maximum rereadings reaches seven times. When we restrict the analysis of time spent on question to questions read only once we indeed find that web and SAQ modes have significantly shorter question reading time than interviewer reading time in face-to-face mode (Table 1, Model 3).

This is an interesting finding that highlights three points: first, questions in web and SAQ modes are not read only once, but are frequently reread after glancing at response options; second, the total time spent on question wording is not drastically different when interviewer reads a question out loud to when a respondent reads a question (including rereadings) in self-completion modes; and the mode difference in total time per question is not explained or explained only in part by the difference in time spent on question wording. Thus, we expect some difference across modes in response options readings.

2 Attention to response options

Response options were presented visually in each mode. In face-to-face mode the interviewer was instructed to not read the response options out loud but to let respondents read them from show cards. Each question had 4 response options presented vertically. Response options were similar, but varied in wording such that only 3 of 12 questions had the same response scale as the preceding question. Because the presentation of the scale is visual in each mode one may expect many similarities in reading the scale across modes. Yet looking at the overall time spent on response options we find that reading them in a face-to-face mode takes more than twice as long as in SAQ mode, and 1.6 times as long (marginally significant) as in web mode (Table 2, Model1).

Table 2. Mode effect on time for response options for all questions and questions with gazes at the bottom options

	Model 1, time for options			Model 2, time for options in q-ns with looks at bottom options		
	Coef.	Linear. SE	p-value	Coef.	Linear. SE	p-value
intercept	4059	695	<0.001	4158	694	<0.001
web (PAPI)	-1530	820	0.09	-1494	873	0.11
SAQ (PAPI)	-2574	714	0.003	-1896	703	0.02
N questions	161			128		
N respondents	14			14		
Model Fit	F(2,12)=7.81, p=0.007			F(2,12)=3.52, p<0.06		

* Time is presented in milliseconds

Visual presentation of response scale is known to be prone to primacy effect (e.g. Galesic et al. 2008), when respondents spend longer time reading top options and may skip looking at the options presented at the bottom of the scale. Previous research has found consistent primacy effect in self-completion modes and in face-to-face mode with show cards, whether read aloud by an interviewer or not (Schwartz et al., 1991). While one may expect similarity in primacy effect across modes when response options are presented only visually in all modes, we find that self-completion modes have over 5 times higher odds of not looking at the bottom 2 response options in comparison to face-to-face mode (p=0.07 for web vs. face-to-face, and p=0.007 for SAQ vs. face-to-face comparisons). And even if we restrict our analysis to questions when the bottom 2 response options are viewed, looking at response options still takes longer in face-to-face mode than in SAQ mode (Table 2, Model 2). So, even if response scale is presented only visually in face-to-face mode, there is some evidence of differences in how respondents view this scale in the presence of an interviewer in comparison to self-completion modes.

3 Comparing attention to question wording versus response options

So far we found that respondents spend less time on question overall, question wording and response scales in self-completion modes than in face-to-face mode. But even if everything is faster in self-completion mode, it is of interest to examine whether there is difference across modes in the proportion of time spent on the question wording rather than response options. Do respondents take longer time for a question wording and rush through response options in one mode but not in other modes? We find that on average respondents take between 1.05 and 1.5 times longer reading response options than listening / reading question wording across modes, but we find no evidence that this ratio is significantly different between modes. The effect of an interviewer reading the question out loud is interesting: this prolongs the whole process, listening to the question and reading response options from a show card, but the ratio of question wording latency to response options latency stays comparable to that in self-completion modes.

Nevertheless, there is an important difference between self-completion and face-to-face modes: while in self-completion mode a respondent cannot view question wording simultaneously with response options, in face-to-face mode respondents can read response options while an interviewer is reading a question out loud. For consistency, interviewers in our experiment instructed respondents to turn to the relevant show card before reading the question. Something that has not been studied before is whether respondents read response options from a show card when an interviewer is reading a question out loud. In our experiment we find that over a third (37%) of the time that an interviewer is reading a question out loud a respondent is simultaneously looking at response options. But this varies across questions with 95% of questions having looks at the response options between 26% and 50% of the time an interviewer is reading a question.

But does this mean that response options are only read when an interviewer is reading the question out loud? We find that on average around half of the time spent on response options (48%, se=2%) is during the time an interviewer is reading a question out loud. This suggests that in a face-to-face mode at least some of the time respondents perceive question wording simultaneously with response options. It is unclear whether this may or may not be optimal for the perception of a question and a scale, given that simultaneous visual and aural perception may be natural to humans. One may think that this potentially may lead to differences in question and response options perception across modes, given that some response options are read before the question is completely heard in a face-to-face mode. But recall that self-completion modes have a number of rereadings of question wording in-between looks at response options, suggesting a possibility of similar simultaneous – or at least interleaved - processing of question wording and response scale in self-completion modes.

4 Order of response option readings

Previous research has found that order of questions in the questionnaire and order of response options in a response scale may influence survey responses. This is often attributed to context effects, or primacy and recency effects (Schwartz et al., 1991). An additional reason why modes may affect responses may be that response options are read in a different order in different modes, even if the response options are presented in the same way. Recall, response options in our study are presented only visually in all modes and there is no reason to expect that the order of their readings differ across modes.

For each question we coded whether the first look at the top response option is before the first look at the next response option (options 1,2), whether the first look at the second top option is before the first look at the third top option (options 2,3), and similarly for options 3 and 4 (options 3,4). If a response option was not viewed, the order measure involving this option was set to missing. Note, for some of the questions we could not differentiate between glances at a specific option. These glances are excluded from this analysis.

Overall, we find similarity in sequential viewings for adjacent response options only for the last two categories (Table 3). In each mode the bottom two response options were viewed in sequential order in 2/3 of questions which received viewings for these options. The difference is larger for the first and middle two categories. In face-to-face mode the top option was viewed before the second top option only in 48% of questions, compared to 76% in Web mode and 94% in SAQ mode. Less dramatic difference in the same direction is observed for the sequence of first viewing of options 2 and 3.

Table 3. Percent of questions with sequential viewings of adjacent response options

	PAPI	Web	SAQ	Total
options 1,2	48%	76%	94%	67%
options 2,3	52%	69%	75%	62%
options 3,4	72%	67%	71%	69%

We next counted how many times response options were viewed for the first time in reverse order for each question, and compared these counts across modes. We find that for some questions all options, which were viewed, were first viewed in sequential order, but this proportion varies drastically across modes: there are only 20% of such questions in face-to-face mode compared to 44% in web mode and 83% in SAQ mode (Table 4). Interestingly, only face-to-face mode had questions (8%) for which all response options were read in reverse order, from bottom to top. It is clear that tendency to read response options sequentially is higher in self-completion modes than in face-to-face mode: around 88% of questions in web and 96% in SAQ had one or fewer reverse viewings.

Table 4. Percent of questions with reversed viewings of adjacent response options

	PAPI	Web	SAQ	Total
0	20%	44%	83%	41%
1	52%	44%	13%	42%
2	20%	12%	4%	14%
3	8%	0%	0%	3%

* 0 means that all options viewed were viewed for the first time in sequential order; 1 indicates that the adjacent options were viewed in reversed order once within a question; 2 – twice; and 3 – the response scale was viewed for the first time from bottom to top

Overall, our preliminary results suggest that even if response options are presented only visually in face-to-face mode, respondents often do not read them sequentially from top to bottom. This is different from web and SAQ modes for which sequential reading of response options is dominant.

Discussion

The current paper demonstrates successful use of a real-world eye tracker in a survey methods context, enabling the study of visual attention in a natural setting for such survey modes as SAQ and face-to-face with show cards, in addition to the previously explored web mode. With minor restrictions, the eye-tracker can track eye-movements wherever a respondent may look with no limitations to head movements or where a respondent may place the paper questionnaire or show cards. For studying eye-movements in a web mode, the real-world eye-tracker has the important limitation of requiring additional manual coding for time span of eye-fixations. It is nevertheless a useful tool when researcher is interested in gazes away from the screen, like at the mouse or the key board, for example for studying web responding among non PC users. While the real-world eye-tracker is a good tool for studying visual attention in each mode separately, it has potential for providing standardized comparable information on visual attention for mode comparison studies using the above three modes.

Additionally, our study provides empirical example of how eye-tracking information can be used for studying survey response process across modes. Specifically, we were able to compare gaze latency when viewing question wording or response options, as well as reading order of response options across modes. We observed some of the unique visual behaviour of respondents, previously not discussed in the survey literature. For example, while one may expect that aural reading of a question should take longer time than visual reading, we find that this is true only for the questions read once. Nevertheless, most questions in self-completion modes receive one or more rereadings, which leads to the overall time taken per question wording being similar across modes, on average. We also find similarity in the ratio of time taken for response options to that for question wording across modes.

A difference arises in how response options are read. Even though response options were presented visually in all modes, we found substantial differences in reading styles of response options between face-to-face mode and self-completion modes. For example, it is more common to read bottom response options in face-to-face mode than in self-completion modes. The reading order of response options varies much across modes: sequential reading of response options was found dominant in SAQ mode followed by web mode, and rather uncommon in face-to-face mode.

Our empirical results also showed evidence of simultaneous consideration of response options and question wording in all three modes: in face-to-face mode it is not uncommon to read response options while an interviewer is reading question wording out loud; and in self-completion modes many questions receive rereadings between glances at the response options. It is unclear to which extent this may influence perception of the question and scale options. But it is clear that when designing questions we cannot assume their sequential reading, starting with question wording, and following with each response options the way they are presented in a questionnaire or a show card; and that perception order of question and response options is similar across modes.

Given the fast move of survey data collection to mixed modes, we need to better understand the potential influence of modes on responses. In a search for mode-resistant questions or for questions comparable across modes clear understanding is required on how each feature of a question may affect perception and attention, and through these, potentially affect responses. The current paper is limited in exploring visual attention without linking it to the choice of responses. We are looking forward to future research studying this.

References

Ashenfelter, Kathleen T., and Todd R. Hughes (2011). Results from Iterative Usability Testing of the American Community Survey (ACS). *Paper presented at European Survey Research Association Conference,* Lausanne, Switzerland

Bilgen, Ipek, and Belli Robert F. (2010) Comparison of Verbal Behaviors between Calendar and Standardized Conventional Questionnaires, *Journal of Official Statistics*. 26(3): 481–505

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health, 27*(3), 281-291.

Burke, Moira, Anthony Hornof, Erik Nilsen, and Nicholas Gorman. 2005. High-cost banner blindness: Ads increase perceived workload, hinder visual search, and are forgotten. *ACM Trans. Comput.-Hum. Interact*. 12, 4, 423-445.

Dillman, Don A., Smyth Jolene D., Christian Leah M (2009). *Internet, Mail, and Mixed-mode Surveys: The Tailored Design Method*. Hoboken, NJ: Wiley

Federal Statistical Office. (2011). *Methods – Approaches – Developments: Information of the German Federal Statistical Office.* Wiesbaden: Federal Statistical Office.

Foulsham, T., Walker, E. & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research.* 51(17): 1920-1931.

Fouse, A., N. Weibel, E. Hutchins and J. Hollan. (2011). "ChronoViz: A system for supporting navigation of time-coded data" In Extended Abstracts of CHI 2011, SIGCHI Conference on Human Factors in Computing Systems, Vancouver, Canada, May 2011.

Freeth, Megan, Tom Foulsham, and Alan Kingstone. (2013). What affects social attention? Social presence, eye contact and autistic traits. PLoS ONE, 8(1).

Galesic, Mirta, Roger Tourangeau, Mick P. Couper, Frederick G. Conrad (2008). Eye-tracking Data: New Insights on Response Order Effects and Other Cognitive Shortcuts in Survey Responding. *Public Opinion Quarterly,* 72(5): 892–913

Galesic, Mirta and Ting Yan. 2011. "Use of Eye Tracking for Studying Survey Response Processes." In *Social and Behavioral Research and the Internet: Advances in Applied Methods and Research Strategies*, eds. Marcel Das, Peter Ester and Lars Kaczmirek. London: Routledge Taylor and Francis Group.

Goldberg D, Williams P. 1991. *A User's Guide to the General Health Questionnaire.* London: NFER-Nelson.

Graesser, Arthur C., Zhiqiang Cai, Max M. Louwerse, and Frances Daniel (2006). Question Understanding Aid (QUAID): A Web Facility that Tests Question Comprehensibility. *Public Opinion Quarterly.* 70:1–20

Groves, Robert M., Fowler, F. J., Couper, Mick P., Lepkowski, James M., Singer, Eleanor, & Tourangeau, Roger (2004). *Survey methodology*. Hoboken, NJ: John Wiley and Sons.

Hayhoe, Mary, and Dana Ballard (2005) Eye movements in natural behaviour. *Trends in Cognitive Sciences*. 9(4): 188-194

Hicks, Wendy D., Brad Edwards, Karen Tourangeau, Brett McBride, Lauren D. Harris-Kojetin, Abigail J. Moss (2010).Using Cari Tools To Understand Measurement Error. *Public Opinion Quarterly,* 74 (5): 985-1003.

Jans, Matt (2009). *Verbal Paradata and Survey Error: Respondent Speech, Voice, and Question-answering Behavior can Predict Income Item Nonresponse*. Unpublished dissertation, Program in Survey Methodology, University of Michigan

Kamoen, Naomi, Bregje Holleman, Pim Mak, and Ted Sanders (2011). Agree or Disagree? Cognitive Processes in Answering Contrastive Survey Questions. *Paper presented at European Survey Research Association Conference,* Lausanne, Switzerland

Menold, Natalja, Lars Kaczmirek, and Timo Lenzner (2011). Cognitive Processes in Answering Questions in Web surveys: How do Respondents Attend to Verbal Labels in Rating Scales? *Paper presented at European Survey Research Association Conference,* Lausanne, Switzerland

Land M F, Lee D N, (1994). Where we look when we steer. *Nature (London)* 369, 742-744

Land, M.F., and P. McLeod (2000). From eye movements to actions: how batsmen hit the ball. *Nat. Neurosci.* 3: 1340–1345.

Land,M.F.,Mennie,N., and Rusted,J.(1999).The roles of vision and eye movements in the control of activities of daily living. *Perception*. 28(11): 1311-1328.

de Leeuw, E. D. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics, 21*(2), 233-255.

Lenzner, Timo, Lars Kaczmirek, and Mirta Galesic (2011a). Seeing Through the Eyes of the Respondent: An Eye-tracking Study on Survey Question Comprehension. *International Journal of Public Opinion Research*: 23(3): 361-373.

Lenzner, Timo, Lars Kaczmirek, Mirta Galesic, and Mona Merkert (2011b). Left feels right! The optimal position of answer boxes in Web surveys. *Paper presented at European Survey Research Association Conference,* Lausanne, Switzerland

Ongena, Yfke (2010). *Interviewer and Respondent Interaction in Survey Interviews: Empirical evidence from behaviour coding studies and question wording experiments*. LAP Lambert Academic Publishing

Potaka, Lyn. 2007. What the Eye doesn't see: A Feasibility Study to Evaluate Eye-Tracking Technology as a Tool for Paper-Based Questionnaire Development. *Proceedings from Quest 2007 Statistics Canada*, Ottawa, Canada: 162-171.

Schwarz, N., Strack, F., Hippler, H., & Bishop, G. (1991). The impact of administration mode on response effects in survey measurement. *Applied Cognitive Psychology, 5*, 193-212.

Rayner K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology,* 62 (8), 1457-1506

Redline, Cleo D., and Christopher P. Lankford (2001). *Eye-Movement Analysis: A New Tool for Evaluating the Design of Visually Administered Instruments (Paper and Web)*. Proceedings of the Survey Research Methods Section of the American Statistical Association, available at www.amstat.org/Sections/Srms/Proceedings/y2001/Proceed/00248.pdf.

Redline, D. Cleo, and Don A. Dillman, (2002). The Influence of Alternative Visual Designs on Respondents' Performance with Branching Instructions in Self-Administered Questionnaires. In R. M. Groves, D. A. Dillman, J. L. Eltinge & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 179-196). New York: John Wiley and Sons, Inc.

Reichle, E., D., Reineberg, A.E. and Schooler, J. K. (2010). Eye Movements During Mindless Reading. *Psychological Science*, 21: 1300-1310

Romano, Jennifer C. and Jennifer M. Chen. 2011. "A Usability and Eye-tracking Evaluation of Four Versions of the Online National Survey of College Graduates (NSCG): Iteration 2." *Study Series: Survey Methodology 2011-01.* Washington D.C.: U.S. Census Bureau.

Schwarz, Norbert. *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation.* Mahwah: New Jersey: Lawrence Erlbaum Associates, Publishers.

Tourangeau, R., Rips, L. J., & Rasinski, K. A. (2000). *The psychology of survey response*. Cambridge, United Kingdom: Cambridge University Press.

Underwood, G., & Foulsham, T. (2006). Visual saliency and semantic incongruency influence eye movements when inspecting pictures. *QuarterlyJournal of Experimental Psychology:* 59 (11), 1931 - 1949.

Willis, Gordon B. (2005). *Cognitive Interviewing: A Tool for Improving Questionnaire Design.* Thousand Oaks, CA: Sage Publications, Inc.