Self-report questionnaires scrutinised : do eye movements reveal individual differences in cognitive processes while completing a questionnaire?

Self-report questionnaires scrutinised: Do eye movements reveal individual differences in cognitive processes while completing a questionnaire?

Margot Chauliac[a], Leen Catrysse[b], David Gijbels[a], Sven De Maeyer[a] and Vincent Donche[a]aUniversity of Antwerp, Antwerp, BelgiumbOpen University

Margot Chauliac is a doctor in Educational Sciences. Her main research interests are on educational measurement through self-report questionnaires and eye movement registration. Currently she works as apolicy advisor at the Antwerp Doctoral School at the University of Antwerp. Here she focuses on international research projects.
Leen Catrysse currently works as an assistant professor at the group Online Learning and Instruction at the Faculty of Educational Sciences at the Open University in the Netherlands. Leen her research focuses on student learning in higher education and measuring learning and reading processes with online measures such as eye movement registration.
David Gijbels is professor of learning and instruction at the Department of Education and Training Sciences at the University of Antwerp. His research is situated in the research group Edubron and the Antwerp Social Lab and focuses on student learning in (higher) education and in the workplace.
Sven De Maeyer is a full professor at the Department of Education and Training Sciences at the University of Antwerp. His research focuses on the measurement of (learning) processes, (formative) assessment and rating in education (with an emphasis on the merits of comparative judgment) and (Bayesian) mixed-effects modelling.
Vincent Donche is a professor at the Faculty of Social Sciences, Department of Training and Education Sciences, University of Antwerp. Within the research group EduBROn, he conducts research in the domains of learning and instruction, higher education and educational measurement.

CONTACT Margot Chauliac margot.chauliac@uantwerpen.be University of Antwerp, Middelheimlaan 1, Antwerpen 2020, Belgium

KEYWORDS
Eye-tracking
working memory capacity
processing effort
cognitive processes
mixed effect models

Introduction

Answering self-report questionnaires is assumed to be an easy task. After reading the question stem of an item of a questionnaire, the respondent answers it and moves to the next item until the questionnaire has been completed. However, the credence that this is an easy process is merely an illusion. Completing a questionnaire is a complex task that involves many cognitive processes (Karabenick et al., 2007; Tourangeau, 1984).

Despite the long history of survey research, it was only in the 1980s that the cognitive aspect of survey methodology (CASM) arose (Miller, 2014). Prior to the CASM-movement, focus in this area solely was on the outcomes of the questionnaire. However, attention has shifted to the cognitive processes involved in completing questionnaires (Fowler, 2014; Willis &

Miller, 2011). The CASM-movement established the idea that individual processes must be understood to assess the validity of a questionnaire and the possible sources of error (Schwarz, 2007). The key assumption is that respondents' cognitive processes drive survey responses, so understanding cognition is central for understanding responses and reducing errors (Willis, 2004).

The basic insight of the CASM-movement is that respondents give inaccurate answers for a variety of reasons. Research shows that (1) respondents' understanding of the item questioned can be doubtful, (2) they might not remember all relevant information to form a judgment, (3) their judgment can be flawed or (4) their estimation strategies might be imprecise. Respondents may also (5) have trouble pinpointing their internal judgment on a possible response category or (6) change their answers deceitfully before reporting them (Tourangeau, 2003). However, despite these insights into the problems that can arise when completing self-report questionnaires, there still lies a veil over what exactly is going on in the minds of the respondents during the completion of a questionnaire. By distributing a self-report questionnaire, researchers still very much rely on the respondents' perceptions of their cognitive processes. To unobtrusively track the cognitive process at play, eye tracking has proven to be a useful means.

Eye-tracking is the process of recording eye movements as a person completes a task. It has recently been introduced into the field of survey methodology to study cognitive processes that occur when a person answers survey questions. Eye-tracking provides detailed information on the cognitive processes taking place when respondents complete questionnaires, and it shows in detail how processing a question stem influences the processing of the response categories. The relationship between eye movements and cognitive processing is based on two assumptions: the immediacy assumption and the eye-mind assumption. The first assumption postulates that the visual stimulus on which the eyes fixate are processed immediately. The later states that when the stimuli are fixated, they are also mentally processed. Taken together, both presumptions suggest that eye movement provide direct information about what is processed on the one hand, and the amount of cognitive effort involved on the other hand (Just & Carpenter, 1980).

Early studies that used eye-tracking in CASM research focused mostly on the object of the self-report questionnaire, not on the respondent completing it. In addition, those studies were mostly experimental research that focused on manipulating aspects of a questionnaire that could lead to difficulties in processing, such as different response formats (Lenzner et al., 2014), the order of responses (Galesic et al., 2008) and the wording of questions (Graesser et al., 2006; Lenzner et al., 2011). By investigating the potential burden the questions might place on respondents, they focused on the constraints of the survey (Höhne, 2019; Lenzner et al., 2011; Menold et al., 2014; Neuert & Lenzner, 2016). In contrast to that approach, which mainly aimed to investigate different question forms, little is known about the actual process of completing a questionnaire through direct observational measures.

Studies have also shown that individual differences, such as effort and ability, have a large impact on cognitive processes and eye movements during information processing (Catrysse et al., 2018; Krosnick, 1991). Therefore, this study aimed to consider both respondent ability and respondent effort which might lead to differences in processing behaviour.

Theoretical framework

Research in survey methodology starts from the assumption that responding to survey items requires several repetitive steps of information processing (Schwarz, 1990; Sudman et al., 1996; Tourangeau & Hanover, 2018). The process begins by understanding the item, and it continues with retrieving all relevant information from memory. judgment and estimation follow, and they are succeeded by matching the internally generated response to the response categories provided by the questionnaire (Tourangeau, 1984; Tourangeau & Bradburn, 2010; Tourangeau & Hanover, 2018). One can ony expect reliable and valid answers to the items questioned when respondents thoroughly conduct the abovementioned cognitive steps (Krosnick, 1991; Krosnick & Alwin, 1987; Vannette & Krosnick, 2014). However, this can be a very demanding process. Not only must respondents comprehend the item; they must understand them in the same way as the researcher intended (Collins, 2003; Fowler, 2014; Fowler & Cosenza, 2008). Several things can go wrong. Krosnick (1991) discerned three influencing factors that could influence the cognitive processes while completing a questionnaire: task difficulty, respondent ability, and respondent effort. Considering that earlier research was centred on how difficulties in the content of a questionnaire affected processing behaviour, we focused on the latter two: how do respondent's ability and effort influence the process of completion?

Respondent ability

The proficiency to perform mental operations, also termed cognitive sophistication, has been defined as the collection of abilities needed to retrieve information from memory and then integrate that information in the respondent's judgements (Krosnick, 1991; Toplak et al., 2014). The degree of cognitive sophistication is determined by congenital factors and by earlier experiences of the respondent (Schuman & Presser, 1996). One of those innate components is the ability to store and retrieve information from memory while performing a complex task. This is called working memory capacity (Baddeley & Hitch, 1974). There is broad acknowledgement that working memory capacity is limited (Ericsson & Kintsch, 1995; Just & Carpenter, 1992) so respondents might not give all response categories as much attention as needed (Krosnick, 1991). Given the introspection needed, completing self-report questionnaires can be very strenuous for respondents, causing extra pressure on their working memory capacity. Limited working memory capacity can influence the ease by which the questionnaire is being processed (Kimball, 1973; MacDonald & Christiansen, 2002). When respondents must process a lot of information, it can stress their working memory capacity, possibly overloading it, which might cause a breakdown of their working memory (Lenzner et al., 2010) leading to a response behaviour whereby respondents possibly take cognitive shortcuts (Galesic & Bosnjak, 2009). Therefore, working memory capacity might influence the degree to which respondents process the questions and response categories, ultimately affecting the processing of the questionnaire (Gathercole & Alloway, 2013; Krosnick, 1991).

Respondent effort

Another factor potentially influencing respondents' cognitive processing is their investment in the questionnaire. At the beginning of a questionnaire, a survey-taker is more prone to invest in completing the survey (Bogen, 1996). However, as the questionnaire proceeds and more questions follow, the likeliness that a respondent completes a survey with full effort declines. When this decline in effort precisely arise is dependant of multiple factors. Respondents may become fatigued, disinterested, or distracted by external factors. Several studies have shown a significant effect of questionnaire length and response rate. Longer questionnaires are not associated only with lower response rates; studies have also shown that as the questionnaire proceeds, questions are processed differently (C. E. Neuert, 2021; Galesic & Bosnjak, 2009; Helgeson & Ursic, 1994). This might lead to a more uniform answering pattern (Herzog & Bachman, 1981) and an overall decline in the quality of data (James & Bolstein, 1990).

In addition, it is important to look at this process across time . Research has shown that some respondents become distracted or disinterested as they complete a questionnaire (Bogen, 1996; James & Bolstein, 1990), so it is likely that variance in reading and response behaviour also are present across different stages of completing a questionnaire.

Eye-tracking as an illuminator during self-report questionnaires completion

Eye-tracking has a long tradition in studying the cognitive processes that occur during reading and other information processing tasks (Duchowski, 2007; C. Neuert, 2016; Rayner, 1998). By recording the position of a person's gaze, one can gather rich and precise data on the mechanisms of cognitive processing (Beesley et al., 2019; Galesic et al., 2008). Two commonly used measures in reading research are the first- and second-pass fixations. The first-pass fixation duration indicates the time spent reading a specific area of interest when visited for the first time. The second-passfixation duration refers to the total time spent rereading this area after the first-pass reading was terminated (Hyönä et al., 2003). The first-pass fixation duration indicates the early stages of processing without strategic or conscious behaviour. However, the second-pass fixation duration represents more strategic and conscious processing. Therefore, a longer second-pass reading can be seen as a strategic attempt to resolve problems with comprehension (Ariasi et al., 2017; Hyönä & Lorch, 2004; Hyönä et al., 2003).

Cognitive processing

A recent study focused on understanding the process of survey completion by analysing the processing of the question stem and the response categories in an integrated way (Chauliac et al., 2020). However, given the different characteristics of both parts of the questionnaire, and the different cognitive processes that may take place (reading, judging and scoring items) specific attention to the question stem, as well as the response categories, is warranted (Tourangeau & Bradburn, 2010; Tourangeau & Hanover, 2018). Respondents may, for example, show hesitation while reading and processing the question stem, but not while they read and evaluate the different response categories or vice versa. Gaining a better understanding of the relation between the reading and response behaviour of the respondent is paramount, since it could lead to increased insight into the quality of the process when respondents complete self-report questionnaires.

Present study

Existing research has used eye-tracking to detect problems with the object of the questionnaire itself. So, little is known about the role of the individual respondent in terms of (1) the cognitive processes involved in completing the questionnaire and (2) how respondent ability and effort influence processing behaviour. Our study aims at extending the current research on the processing of self-report questionnaires. By using eye-tracking we can unobtrusively map the cognitive processing behaviour of respondents as they complete self-report questionnaires. In our study, we take an extra step by not only looking at the item in a questionnaire as a whole, but by treating the question stem and the response categories as two separate entities. The term *item* is used here as acollective term for both the question stem and the response categories (Figure 1). In addition, research attention is given to the question stem, the response categories within each item, and the relationship between the two.

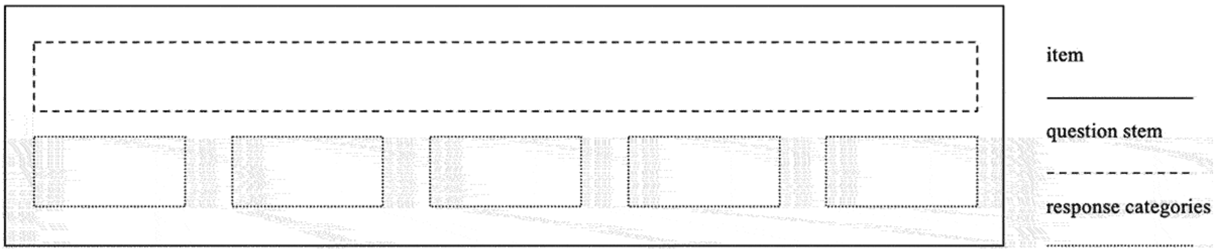

Figure 1.Visualisation terminology item, question stem and response categories.

The following research questions were central to this study.

1. How does a respondent's ability influence their processing of (1) an item, (2) its question stem and (3) its response categories?
2. How does a respondent's effort influence their processing of (1) an item, (2) its question stem and (3) its response categories?
3. How does processing the question stem of a questionnaire influence the processing of the response categories?

Method

Participants

Ninety-two students from a faculty of social sciences were recruited during regular lectures to voluntarily take part in our study. They were informed that this study is part of a larger project about learning from texts and completing self-report questionnaires, whereby eye movements are recorded to gain insights into the processing behaviour of participants. All participants had normal or corrected-to-normal vision, and they were native speakers of Dutch. Because of issues that are common in eye-tracking research (e.g. problems with calibrating the eye tracker, and a lack of visible responses to the survey questions; Holmqvist et al., 2011), we had to exclude data of 17 respondents. After that data cleaning, data from 75 participants were included in the statistical analyses. To thank the students for their participation in the study, they received two cinema tickets.

Instruments and procedure

Questionnaire

After reading an expository text, respondents completed a task-specific self-report questionnaire to measure their cognitive processing strategies. For this, the scales about cognitive processing strategies of the ILS-SV questionnaire were used (Donche & Van Petegem, 2008; Vermunt & Donche, 2017). The questionnaire had 16 items, covering four scales: (1) relating and structuring, (2) memorising, (3) critical processing and (4) analysing (Table 1). The aim was to map how participants cognitively process information while reading a text. Students were asked to read aquestion stem, select the response category that matched their behaviour best, and state their answer out loud. Response categories ranged from 1 = ‘I rarely or never do this’ to 5 = ‘I almost always do this’. All items had to be answered consecutively and answers could not be changed later.

Table 1.ILS-SV scales, number of items and item examples.

Scale	Items	Item example
Relating and structuring	4	I compare conclusions from different teaching modules with each other.
Critical processing	4	I try to understand the interpretations of experts in a critical way.
Analysing	4	I study each course book chapter point by point and look into each piece separately.
Memorising	4	I learn definitions by heart and as literally as possible.

Ability

Respondents’ ability was operationalised on the basis of their working memory capacity, measured by employing the Automated Operation Span Task (Aospan; Unsworth et al., 2005). Before the reading of the text and the completion of the questionnaire, participants were required to solve aseries of mathematical operations while trying to retain a set of unrelated letters in their memory. To be sure that participants were not focusing on only remembering the letters, an 85% accuracy criterion was imposed for solving the mathematical problems (Unsworth et al., 2005). Aospan provides two scores, an absolute and a partial credit score. Since partial credit scoring is preferred over the absolute credit scoring, we used the latter (Conway et al., 2005). The mean partial score for all respondents was 59.533 (*SD* = 10.687). The score for this working memory capacity test was normally distributed, and standardised scores were used for further analysis.

Effort

Research shows that as the questionnaire proceeds, questions are processed differently, which might decrease data quality (Baer et al., 1997; Meade & Craig, 2012). To this finding, we concretised respondents’ proceedings through a questionnaire as an important indicator of their invested effort. To gain more insight into this aspect of progress in a questionnaire, we divided our questionnaire into four phases. As our questionnaire had 16 items, each phase had four items which repeatedly questioned the four different cognitive processing strategies. Because the survey questions measure four concepts that were presented in a mixed way, we could ensure that one item per concept was present in each phase. We did not rotate the question blocks, all items were administered in the same order for all participants.

Processing behaviour

A Tobii Pro X3-120 eye-tracker was used to measure respondents’ processing behaviour. This eye-tracker alternates between bright and dark pupil eye-tracking in a predefined, systematic way. Tobii Pro X3-120 has a sampling frequency of 120

Hz (binocularly), which makes it possible to look more closely at the fixation durations. The eye tracker was secured to a 17.3-inch monitor with are solution of 1920 × 1080 pixels. Every participant sat about 60 cm from the monitor screen. To minimise the influence of student movement, we used a chin rest. Tobii Technology (Stockholm, Sweden) reported a gaze accuracy of 0.4°, gaze precision of 0.24° and a total system latency of less than 11 milliseconds for this eye tracker. The eye movements were recorded with Tobii-Studio(3.4.8) software.

Analysis

Eye-tracking measures

To map the respondents' gaze patterns, we used the Tobii fixation filter for fixation identification. The filter used a classification algorithm proposed by Olsson (2007), which had a velocity threshold of 35 pixels per window and a distance threshold of 35 pixels (Olsen, 2012). We considered the item, the question stem and the response categories in the questionnaire as distinct areas of interest (AOI). For each AOI, the total fixation was calculated. To analyse the question stem and the response categories, first- and second-pass reading times were calculated as well (Hyönä et al., 2003). Since a respondent can switch between the question stem and response categories, these measures can be more informative about the level of these AOI's than the total fixation duration alone.

To control for the length of AOI's, the duration measures were normalised by calculating amilliseconds-per-character measure (Ariasi et al., 2017; Catrysse et al., 2016; Yeari et al., 2016). In addition, we logarithmically transformed them, using log1p, by which zeros remain zeros after transformation (Holmqvist et al., 2011; Lo & Andrews, 2015).

Mixed effect modelling

The eye movement data were analysed with linear mixed-effects models (LMM) with the lme4 package (Bates et al., 2015) in R (R Core Team, 2014). Mixed-effects models are statistical models that incorporate random and fixed effects (Baayen et al., 2008). Respondents and items were considered as crossed random effects (Baayen, 2008; Baayen et al., 2008). By treating students as random effects, we took the variability associated with each student explicitly into account in the analysis (Baayen, 2008). By doing so, we did control for individual differences without trying to further explain them by other fixed effects such as reading speed, motivation, etc. Our analyses were conducted at the item level, and they were based on 1200 data points (75 students each processing 16 items).

To respond to the first two research questions, different models were estimated for the (a) item, (b) question stem and (c) response categories. Regarding the item, the total fixation duration was the dependent variable, and fixed effects were added for the reading phase, working memory capacity and their interaction. We estimated a random intercept model and a random slopes model (allowing the effect of the reading phase to vary between participants), and we compared the models based on the likelihood-ratio test. We used Akaike's Information Criterion (*AIC*) to rank the models based on their plausibility. The lower the AIC, the better the model represented the reality given in the data. The best-fitting model was then used for interpreting the results (see Appendix, Table A for the statistical comparison of the different models). Table 2 shows the best-fittingmodels. To compare the different phases with each other, multiple comparisons of means (Tukey contrasts with Bonferroni correction) were calculated using the multcomp package (Bretz et al., 2010; Appendix, Table B).

Table 2.Overview of reported mixed-effects models.

	Item	Question stem	Response categories
Total fixation duration	random slopes	random intercept	random slopes
First-pass fixation duration	–	random intercept	random intercept
Second-pass fixation duration	–	random intercept	random intercept

Also, to formulate an answer to the third research question, a random intercept model and arandom slopes model were estimated. The latter was the better-fitting model so it was used for interpreting the results (Appendix, Table A).

Results

Processing the item, the question stem and the response categories

The item

Comparison of the two models, a random intercept model and a random slopes model, for the total fixation duration at the item level, showed that the random slopes model had the better fit ($\chi^2(9) = 47.065$ p <.001) (Appendix, Table A). This fit means that there was a difference between respondents regarding the impact of the reading phase on the total fixation duration. As seen in Figure 2, respondents differed more from each other concerning the total fixation duration to process an item in the first phase, and they became more alike when processing items in a later phase.

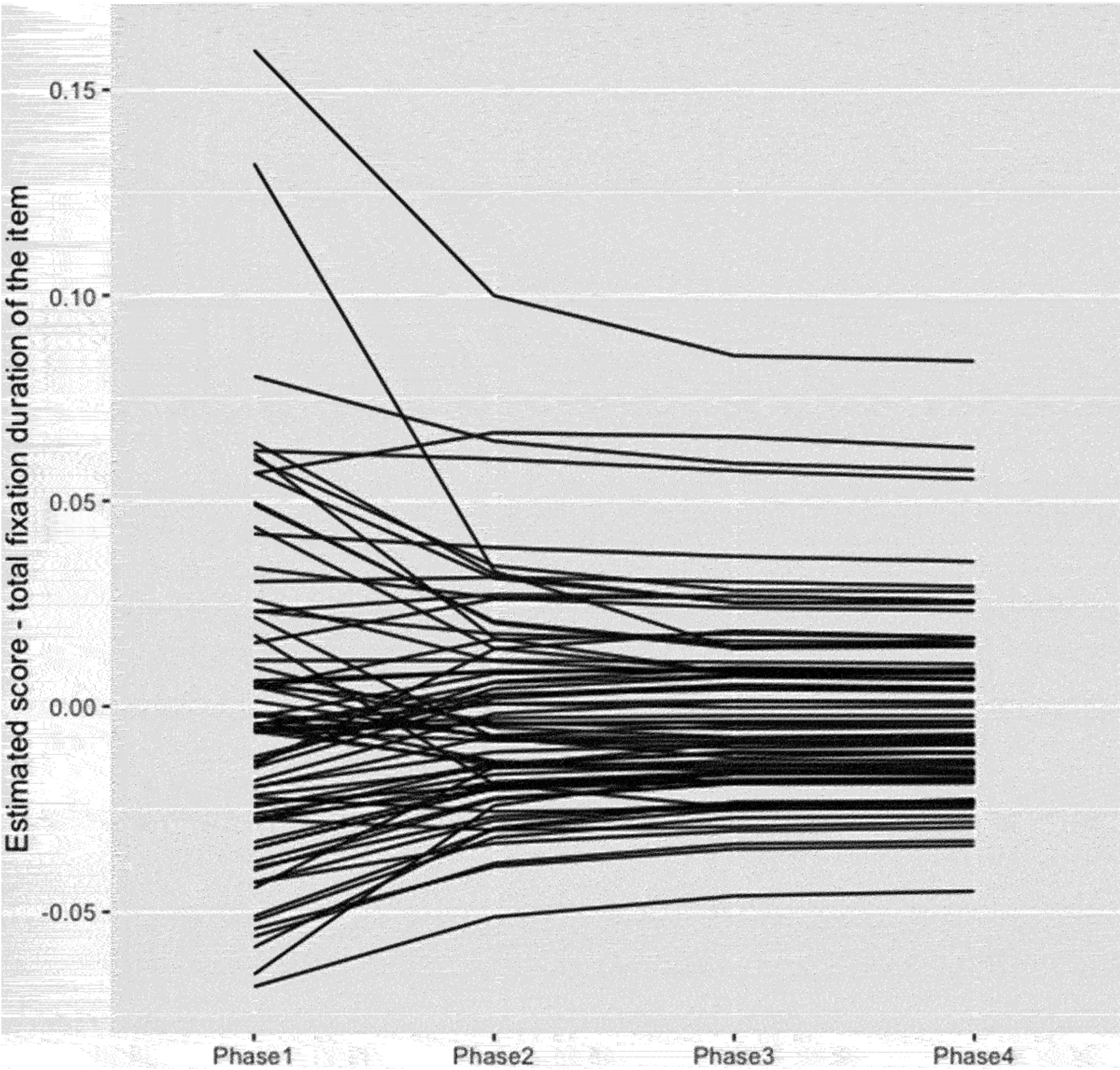

Figure 2.Random slopes model with the effect of phase on the total fixation duration of an item (each line is a respondent).

Results showed a significant effect of phase on how items were processed. Respondents were likely to process an item significantly more slowly during the first phase of a questionnaire, compared to the other phases (Table 3). Additional Tukey post hoc comparisons with Bonferroni correction showed that there were no significant differences between Phases 2 and 3 (p = 1.00), Phases 2 and 4 (p = 1.00) and Phases 3 and 4 (p = 1.00) (Appendix, Table B). This suggests that respondents' effort might deteriorate as they progress through the questionnaire. Phase and working memory capacity do not seem to interact significantly, nor does working memory alone have an effect.

Table 3.Parameter estimates of the random and fixed effects for the random slopes model of the total fixation duration of the item.

Item Total Fixation Duration				
Random effects	Variance		*SD*	

Item Total Fixation Duration				
Respondent level				
Intercept	.002		.046	
Phase 2	<.001		.028	
Phase 3	.001		.033	
Phase 4	.001		.032	
Item level				
Intercept	<.001		.018	
Residual	.002		.044	
Fixed effects	*β*	*SE*	*t*	*pr*(>|*t*|)
Intercept	.163	.011	15.041	**<.001**
WMC	−.007	.006	−1.235	.221
Phase 2	−.050	.014	−3.631	**.003**
Phase 3	−.047	.014	−3.412	**.004**
Phase 4	−.047	.014	−3.410	**.004**
WMC* phase 2	.003	.005	.687	.494
WMC* phase 3	.005	.005	1.029	.307
WMC* phase 4	.002	.005	0.308	.759

Note. Significant values are in bold ($p < 0.05$). Reference category = phase 1.

The question stem

When we compared random intercept and random slopes models for analysing how question stems were processed, the random intercept model appeared to have the better fit for three metrics: the total fixation duration on the question stem ($\chi^2(9) = 8.453$, $p > .05$), the first-pass fixation duration ($\chi^2(9) = 11.895$, $p > .05$) and the second-pass fixation duration ($\chi^2(9) = 5.036$, $p > .05$) (Appendix, Table A). This means that solely considering the processing of the question stem, the effect of the reading phase was not remarkably different from respondent to respondent.

The random intercept model of the total fixation duration of the question stem shows a significant effect of phase on processing behaviour in the second and third phases (Table 4). Additional Tukey post hoc comparisons with the Bonferroni correction showed no differences between Phases 2 and 3 ($p = 1.00$), Phases 2 and 4 ($p = 1.00$) and Phases 3 and 4 ($p = 1.00$) (Appendix, table B). Like the total fixation duration of the item, no significant effect was found for the interaction between working memory capacity and phase. Respondents' ability, as measured by working memory capacity alone, however, did seem to influence processing behaviour. Respondents with a lower working memory capacity needed more time to process a question stem compared to respondents with a higher working memory capacity (Table 3).

Table 4.Parameter estimates of the random and fixed effects for the random intercept model for processing a question stem.

	Question stem											
	Total Fixation Duration				**First-pass fixation duration**				**Second-pass fixation duration**			
Random effects	Variance		*SD*		Variance		*SD*		Variance		*SD*	
Respondent level												
Intercept	<.001		.022		.078		.280		.666		.816	
Item level												
Intercept	<.001		.012		.033		.183		.061		.246	
Residual	.001		.034		.297		.545		1.734		1.317	
Fixed effects	β	*SE*	*t*	*pr(>|t|)*	β	*SE*	*t*	*pr(>|t|)*	β	*SE*	*t*	*pr(>|t|)*
Intercept	.097	.007	14.082	**<.001**	3.765	.102	36.959	**<.001**	2.951	.173	17.099	**<.001**
WMC	−.006	.003	−1.993	**.048**	−.030	.045	−.660	.510	−.192	.121	−1.584	.115
Phase 2	−.022	.009	−2.402	**.033**	.072	.137	.528	.607	−.537	.205	−2.624	**.022**
Phase 3	−.026	.009	−2.874	**.014**	<.001	.137	−.005	.996	−.411	.205	−2.008	.068
Phase 4	−.018	.009	−1.948	.075	.162	.137	1.187	.258	−.563	.205	−2.752	**.018**
WMC* phase 2	.003	.003	.930	.352	.022	.045	.502	.616	.123	.108	1.143	.254
WMC* phase 3	.004	.003	1.464	.144	.030	.045	.668	.504	.267	.108	2.484	**.013**
WMC* phase 4	<.001	.003	.260	.795	.068	.045	1.530	.126	.123	.108	1.141	.254

Note. Significant values are in bold ($p < 0.05$). Reference category = phase 1.

Regarding the first-pass fixation duration processing question stems, Phases 2, 3 and 4 did not differ significantly from Phase 1 (Table 4) nor did additional Tukey post hoc Bonferroni tests show differences between the other phases (Appendix, Table B).

Processing behaviour in the second-pass fixation duration showed a significant effect in Phases 2 and 4. Respondents spent significantly less time rereading the question stem in those phases compared to the first phase (Table 4). Additional Tukey post hoc Bonferroni tests between Phases 2 and 3 (p = 1.00), Phases 2 and 4 (p = 1.00) and Phases 3 and 4 (p = 1.00) showed no differences in processing between those stages (Appendix, Table B). A positive significant interaction effect between working memory capacity and Phase 3 was found (Table 3). Compared to Phases 1, 2 and 4, where a lower working memory capacity resulted in briefer processing, we saw the opposite in the third phase. A higher working memory capacity resulted in a significantly slower second-pass forward processing (Figure 3).

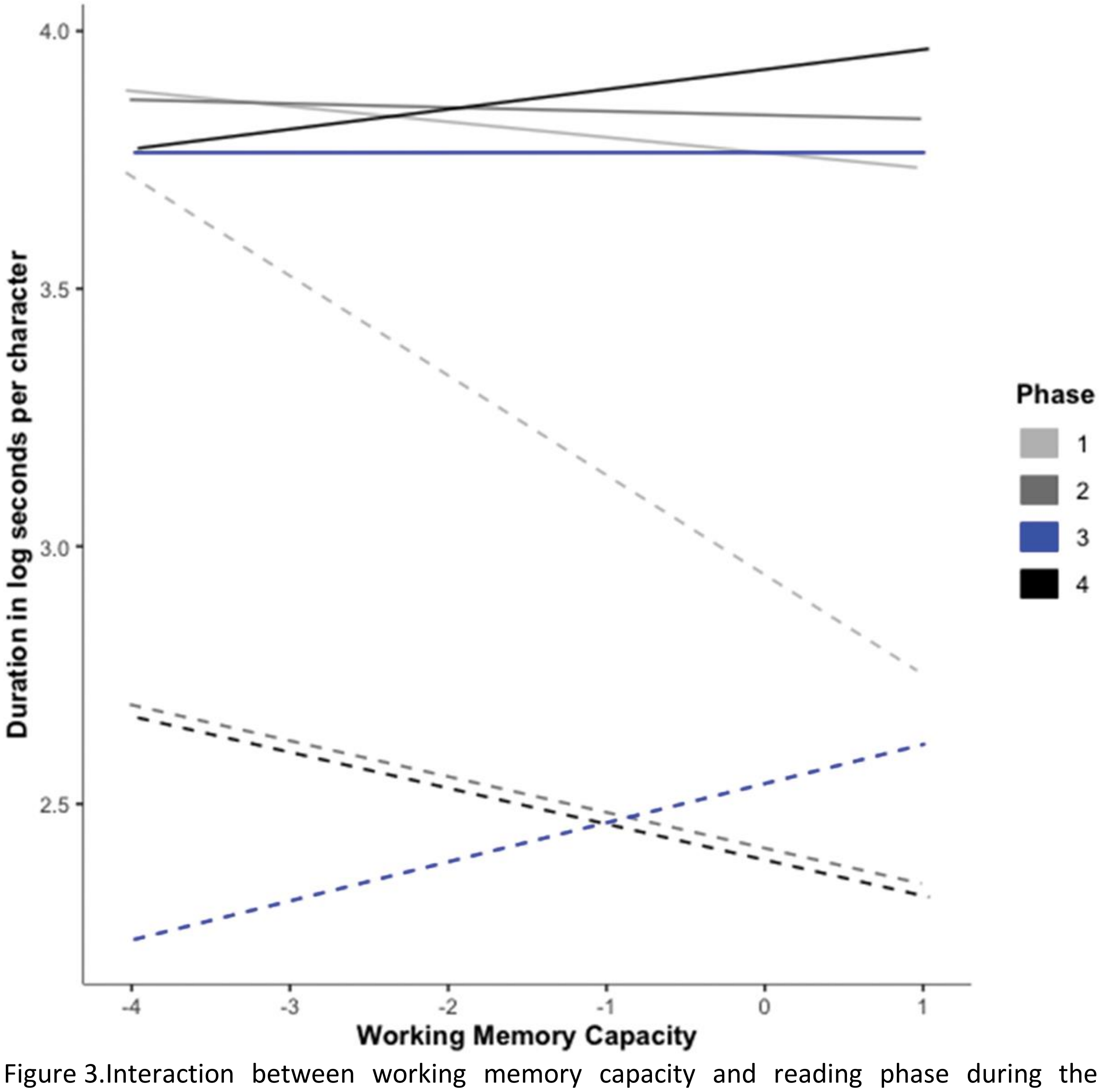

Figure 3.Interaction between working memory capacity and reading phase during the first- andsecond-pass fixation duration of the question. (– = first-pass fixation duration, – = second-pass fixation duration)

The response categories

Comparing both random intercept and random slopes models for the different measures of processing the response categories, we concluded that for the total fixation duration of the response categories, the random slopes model ($\chi^2(9) = 28.315$, $p < .001$) fit better, meaning that there was a difference between respondents regarding the processing of the response categories. In the first phase, we saw a lot of variation between respondents. As the questionnaire continued, respondents become more alike processing the response categories

(Figure 4). Regarding the first ($\chi^2(9) = 5.293$, $p > .05$) and second-pass fixation duration ($\chi^2(9 = 7.664$, $p > .05$) on the response categories, we opted for the random intercept model (Appendix, Table A). This indicated that for processing the response categories, the effect of the reading phase was not uncommonly different across respondents.

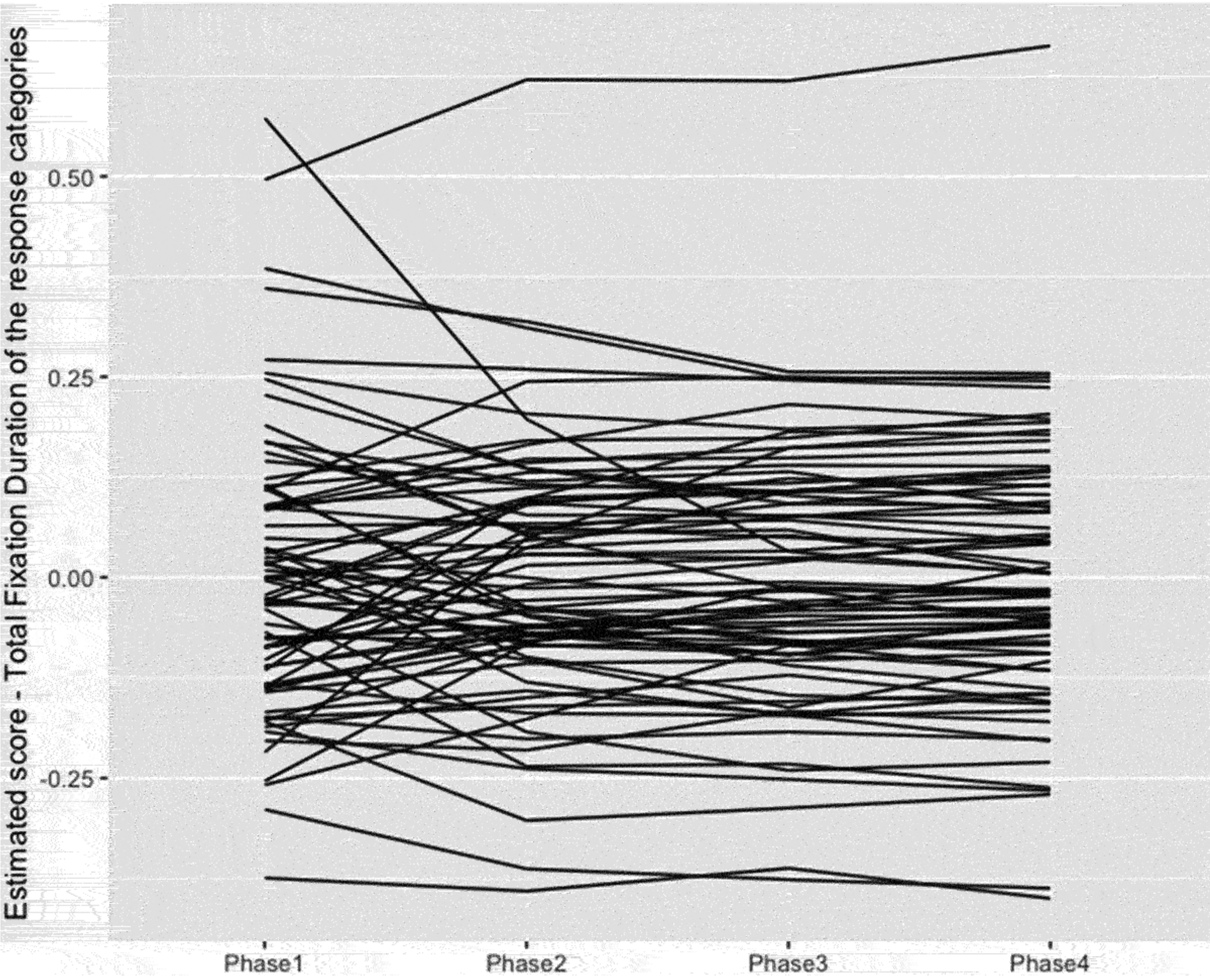

Figure 4.Random slopes model with the effect of phase on the total fixation duration of the response categories (each line is a respondent).

Looking at how respondents processed the response categories on the questionnaire, we saw a significant effect of phase. As respondents proceed through the questionnaire, they spent less time processing the response categories compared to the first phase (Table 5). Additional Tukey post hoc comparisons with the Bonferroni correction showed no significant differences between Phases 2 and 3 (p =1.00), Phases 2 and 4 (p =1.00) and Phases 3 and 4 (p =1.00) (Appendix, Table B). This implies that the total time spent processing the response categories differed significantly for the first phase, where respondents invested more time processing the answers compared to the other phases. Regarding the processing of response categories, phase and working memory capacity did not seem to interact significantly, nor was a significant effect of solely working memory found (Table 5).

Table 5.Parameter estimates of the random and fixed effects for the random slopes and random intercept models for processing response categories.

	Response categories											
	Total Fixation Duration				**First-pass fixation duration**				**Second-pass fixation duration**			
Random effects	**Variance**		***SD***		**Variance**		***SD***		**Variance**		**Variance**	
Respondent level												
Intercept	.042		.206		.041		.203		2.012		1.419	
Phase 2	.021		.145									
Phase 3	.035		.188									
Phase 4	.040		.199									
Item level												
Intercept	.027		.164		.005		.068		.644		.802	
Residual	.076		.276		.327		.572		11.252		3.354	
Fixed effects	β	*SE*	*t*	*pr(>|t|)*	β	*SE*	*t*	*pr(>|t|)*	β	*SE*	*t*	*pr(>|t|)*
Intercept	1.643	.087	18.927	**<.001**	7.357	.053	139.424	**<.001**	5.265	.475	11.094	**<.001**
WMC	−.009	.029	−.299	.766	.068	.041	1.688	.093	−.332	.254	−1.309	.192
Phase 2	−.388	.119	−3.249	**.007**	−.047	.067	−.697	.499	−1.637	.630	−2.599	**.023**
Phase 3	−.421	.120	−3.510	**.004**	−.078	.067	−1.164	.267	−1.662	.630	−2.638	**.022**
Phase 4	−.510	.120	−4.240	**.001**	−.063	.067	−.943	.365	−2.650	.630	−4.207	**.001**
WMC* phase 2	.028	.028	.984	.329	−.041	.047	−.887	.375	.362	.274	1.322	.186
WMC* phase 3	.020	.031	.627	.533	−.054	.047	−1.161	.246	.435	.274	1.589	.112
WMC* phase 4	.019	.032	.585	.560	−.033	.047	−.705	.481	.202	.274	.737	.461

Note. Significant values are in bold ($p<0.05$). Reference category=phase 1

Concerning the first-pass processing, no significant effects of phase were found (Table 4). Respondents spent the same amount of time processing the response categories in all four phases (Appendix, table B). When looking at the second-pass processing, we noticed a significant effect of phase on processing time. Respondents spent significantly less time processing the response categories for a second time in Phases 2, 3 and 4 compared to Phase 1 (Table 4). However, additional Tukey post hoc comparisons with the Bonferroni correction showed that there were no significant differences between Phases 2 and 3 (p =1.00), Phases 2 and 4 (p =.65) and Phases 3 and 4 (p =.70) (Appendix, Table B). For both the first and second-pass processing, no significant interaction effects were found for phase and working memory capacity, nor for solely working memory capacity (Table 4).

The interplay of processing the question stem and the response categories

To answer the third research question, we looked at the first- and second-pass fixation duration of the question stem to see how this influenced the processing of the total fixation duration of the response categories. After comparing the random intercept and random slopes model, the latter showed the better fit ($\chi^2(9)$ = 22.034, p <.001). This indicated that there was a notable difference between respondents based on the processing phase. In the first phase, respondents' processing behaviour was notably more diverse compared to the other phases. In what follows we will, therefore, discuss the random slopes model (Table 6).

Table 6.Parameter estimates of the random and fixed effects for the random slopes model for the influence of the fixation duration of the question stem on the fixation duration of the response categories.

Random effects	**Variance**		***SD***	
Respondent level				
Intercept	.027		.166	
Phase 2	.015		.123	
Phase 3	.026		.162	
Phase 4	.028		.168	
Item level				
Intercept	.018		.135	
Residual	.072		.268	
Fixed effects	β	*SE*	*t*	*pr(>|t|)*
Intercept	1.160	.143	8.120	**<.001**
First-pass fixation duration	.058	.027	2.107	**.035**
Second-pass fixation duration	.090	.012	7.677	**<.001**
Phase 2	−.007	.202	−.037	.971
Phase 3	−.164	.203	−.807	.420
Phase 4	−.210	.200	−1.031	.304
WMC	.010	.025	.420	.676

Random effects	Variance		*SD*	
First-pass fixation*Phase 2	−.066	.041	−1.624	.105
First-pass fixation*Phase 3	−.030	.042	−.711	.478
First-pass fixation*Phase 4	−.029	.040	−.729	.466
Second-pass fixation*Phase 2	−.034	.016	−2.171	**.030**
Second-pass fixation*Phase 3	−.043	.017	−2.564	**.011**
Second-pass fixation*Phase 4	−.060	.017	−3.622	**<.001**
WMC*Phase 2	.012	.026	.474	.637
WMC*Phase 3	−.003	.029	−.100	.921
WMC*Phase 4	<.001	.029	.028	.978

Note. Significant values are in bold ($p < 0.05$). Reference category = phase 1.

Remarkably, compared to the total fixation duration of the question stem and on the response categories, we could not detect a significant effect of phase on the processing of the response categories in general. However, the first- and second-pass fixation duration of reading the question stem had a significant effect on processing the response categories. Respondents who spent more time processing the question stem during the first- and second-pass reading also spent more time processing the response categories. There was no significant interaction between first-pass fixation duration and phase (Figure 5). However, there was significant interaction between the second-pass fixation duration and phase. This implies that, contrary to the first phase, respondents who processed the questions in more depth during the other phases also spent less time reading the response categories. The effect of second-pass fixation duration on total reading time of the response categories varied from phase to phase: the effect was more outspoken in the first phase than it was in all the other phases (Figure 5).

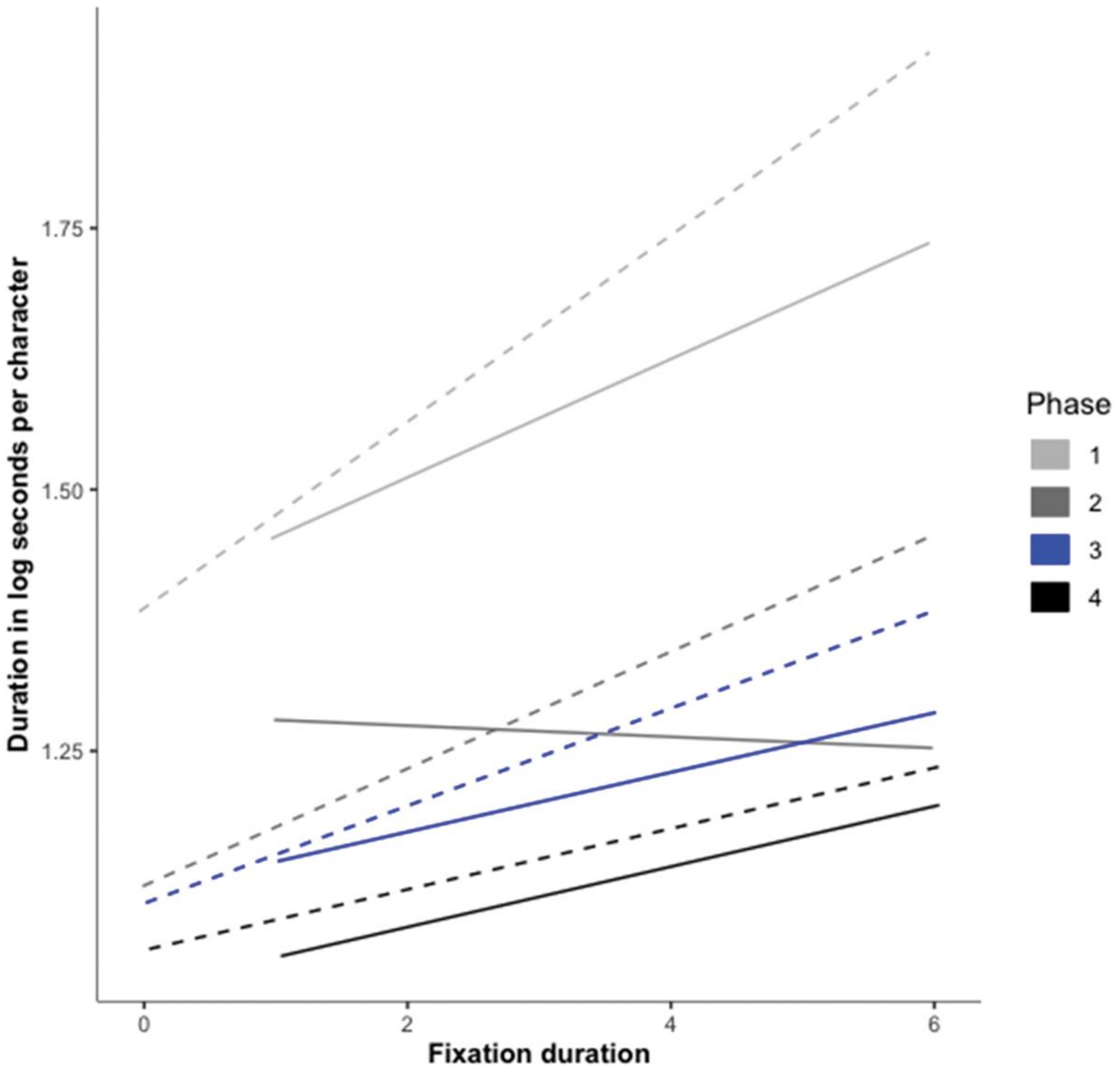

Figure 5.Interaction between the first-pass fixation duration and the total fixation duration and between second-pass fixation duration and the total fixation duration per phase. (– = first-pass fixation duration, – = second-pass fixation duration)

Discussion

Despite the vast majority of research examining the ideal format of items and response categories in questionnaires (Höhne, 2019; Lenzner et al., 2011; Menold et al., 2014; Neuert & Lenzner, 2016), almost no research has focused on the differences in cognitive processes that take place while completing a questionnaire. However, the latter might provide more insight into response patterns or response errors. In our research, we used eye-tracking to gain more insight into the process of completing questionnaires. In the first step, we looked separately at how an item, its question stem and its response categories were being processed. In the second step, we gave more attention to the way that processing the question stem influenced the processing of the response categories. In both steps, we considered how respondents' effort, operationalised by their proceeding through the questionnaire, and their ability, measured as their working memory capacity, influenced this processing.

In our study, we looked at the processing of questionnaires on three different levels: processing the item and processing the question stem and the response categories in that item. Analysis of the processing of the item showed that in the first phase of a questionnaire, respondents showed a more diverse answering pattern and that, as the questionnaire advanced, respondents' processing tended to become more uniform. We found that as the questionnaire proceeded, respondents became faster in their processing of items. This finding confirms previous research that showed that the length of a questionnaire influenced the processing behaviour (Bogen, 1996; Helgeson & Ursic, 1994). To gain more insights into whether this acceleration in processing could be ascribed to respondents becoming accustomed to the questionnaire or whether there was a deterioration of their effort to complete the questionnaire, we did a separate analysis of the question stem and the response categories.

Like the processing of the item, we noted a significant effect of phase on the duration of processing the question stem. Respondents spent significantly more time processing the question stem in the first phase than in the following phases. However, taking a closer look at this processing, and dividing the total fixation duration in a first- and second-pass reading, we noticed this effect of phase only on the second-pass fixation. Thus, even though respondents spent the same amount of effort to process the question stem the first time, we noticed that they spent less time rereading the question as the questionnaire continued. This confirms previous research which showed that as the questionnaire continues, respondents were less prone to invest the necessary effort in the questionnaire so those items were processed differently than the ones near the beginning (Bogen, 1996; Helgeson & Ursic, 1994).

Similar to processing the question stem, proceeding through the questionnaire was shown to influence the processing behaviour on the response categories. In the first phase of the questionnaire, respondents spent significantly more time reading and processing the response categories than in the following phases. This can be attributed to the fact that respondents still had to get used to the response categories and they were not yet fully familiar with the Likert scale on which they were scoring the items. As the questionnaire proceeded, this familiarity increased. This might explain why post-hoc tests showed that the later phases did not differ significantly from each other. Here respondents became more familiar with the response categories and the fact that the same answers recur. Looking more closely, we noticed no significant differences in the first-passfixation, solely when revisiting the response categories. Here again, we see that the answers were processed longer only at the beginning of the questionnaire. The occurrence of this effect only in the second-pass reading might be ascribed to the fact that respondents read the question stem, took the response categories into account and then re-read the question stem so they could give awell-founded answer. Once the response categories were clear, this rereading was of minor importance.

In the next step, we took a closer look at how processing a question stem influenced the processing of the response categories. The analysis showed that when respondents spent more time processing the question stem, they also spent more time processing the response categories. Interaction-effects showed that the effect of second-pass fixation duration on total reading time of the response categories varied from phase to phase: the effect was stronger in the first phase than it was in the other phases.

Previous research stated that the limitations of working memory capacity influenced the ease by which a questionnaire was processed (Kimball, 1973; MacDonald & Christiansen, 2002). However, no effect of working memory capacity was found in the processing of the response categories.

Limitations of the study and directions for further research

Our research demonstrates the propitiousness of using eye-tracking to gain a better understanding of the cognitive processes involved in completing self-report questionnaires. However, despite the favourableness of the technique, we want to indicate some limitations of this study and suggest directions for subsequent research.

As indicated earlier, completing a questionnaire involves different cognitive stages respondents must go through to come to a meaningful result (Tourangeau & Hanover, 2018; Vannette & Krosnick, 2014). Although it was not possible to distinguish between these phases, we tried to gain more insight into parts of this process by treating the question stem and the response categories as two separate AOIs. In our research, we split the process of completing the questionnaire into four distinct phases. We opted for four so that there would be one item of each scale in each phase. Moreover, we also had equally long phases. Results showed that respondents became faster, just as we expected from previous research. However, we would like to emphasise that this was a short questionnaire of only 16 items and that it was taken in a controlled setting. Yet we found this effect. For follow-up research, it would be inventive to remove this controlling factor and work with a longer questionnaire. The effects might be even more pronounced.

Another factor we want to point attention to is the homogeneity of research sample. Our respondents were students in higher education completing a questionnaire on student learning. This is exactly the type of respondents where these questionnaires are commonly distributed, however, this homogenous group also implies that we cannot generalise our results across another population. Concerning follow-up research, it would therefore be interesting to see whether a different research sample, in, perhaps, a different field of research, would confirm the obtained results. Also taking into account other respondent characteristics such as their motivation could be of interest in follow-up research.

A final observation is that the respondents were asked to affirm their answer out loud after every item. Knowing researchers were monitoring their answers might have influenced the natural process of completion. In subsequent research, it would be enlightening to study the cognitive processes without asking respondents to enunciate their answers. Immediately after answering, the next item was projected, so respondents did not have the opportunity to change their answer. By taking this into account in follow-up research, we could learn more about doubts and changes in the answering process.

Conclusion

Notwithstanding some limitations we have noted, we can emphasize the insights into the processing of questionnaires provided by eye movements. The first aim of this study was to investigate whether respondents' effort and ability influenced their processing of the questionnaire. Our results point to the important effect of effort on how both questions and response categories were processed. At the beginning of the questionnaire, respondents were much more likely to spend time processing the presented materials, but as the questionnaire proceeded, this effort declined. Contrary to our expectations, respondents' ability did not have much effect. The working memory capacity of a respondent did not seem to have much of an influence on their processing behaviour. In the following step, we looked at how processing a question stem influenced the processing of the response categories. Here we saw that they were closely intertwined. When respondents spent more time processing a question stem, they did the same for the response categories.

In sum, the use of eye-tracking in this study enabled us to provide a deeper understanding of the individual differences in cognitive processing that take place when respondents complete a questionnaire. Based on the variance indicators, it is clear that not all respondents spend the same amount of time reading and answering items, and that important differences were found regarding the quality of processing. Apparent from these findings is that the processing indicators were different for the first phase of a questionnaire and that the way by which respondents process questions seems to be closely related to their processing of the response categories.

Funding
The research reported in this article was supported by a grant from the "Special Research Fund from the Research Board of the University of Antwerp.

Disclosure statement
No potential conflict of interest was reported by the author(s).

References

Ariasi, N., Hyönä, J., Kaakinen, J., & Mason, L. (2017). An eye-movement analysis of the refutation effect in reading science text. *Journal of Computer Assisted Learning*, *33*(3), 202–221.
https://doi.org/10.1111/jcal.12151

Baayen, R. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.
https://doi.org/10.1017/CBO9780511801686

Baayen, R., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.
https://doi.org/10.1016/j.jml.2007.12.005

Baddeley, A., & Hitch, G. (1974). Working memory Bower, Gordon H.. In *Psychology of learning and motivation* (Vol. 8, pp. 47–89
https://doi.org/10.1016/S0079-7421(08)60452-1
doi:). Elsevier.Baer, R., Ballenger, J., Berry, D. T. R., & Wetter, M. W. (1997). Detection of random responding on the MMPI–A. *Journal of Personality Assessment*, *68*(1), 139–151.
https://doi.org/10.1207/s15327752jpa6801_11

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.
https://doi.org/10.18637/jss.v067.i01

Beesley, T., Pearson, D. S., & Le Pelley, M. (2019). Eye-tracking as a tool for examining cognitive processes. In G. Foster (Ed.), *Biophysical measurement in experimental social science research*. Academic Press 1-30 .
https://doi.org/10.1016/B978-0-12-813092-6.00002-2
Bogen, K. (1996). The effect of questionnaire length on response rates - a review of the literature. *Proceedings of the section on survey research methods* 1020-1025 .Bretz, F., Hothorn, T., & Westfall, P. H. (2010). *Multiple comparisons using R*. CRC Press.
https://doi.org/10.1201/9781420010909
Catrysse, L., Gijbels, D., & Donche, V. (2018). It is not only about the depth of processing: What if *eye* am not interested in the text? *Learning and Instruction*, *58*, 284–294.
https://doi.org/10.1016/j.learninstruc.2018.07.009
Catrysse, L., Gijbels, D., Donche, V., De Maeyer, S., Van den Bossche, P., & Gommers,L. (2016). Mapping processing strategies in learning from expository text: An exploratory eye tracking study followed by a cued recall. *Frontline Learning Research*, *4*(1), 1–16.
https://doi.org/10.14786/flr.v4i1.192
Chauliac, M., Catrysse, L., Gijbels, D., & Donche, V. (2020). It is all in the surv-eye: Can eye tracking data shed light on the internal consistency in self-report questionnaires on cognitive processing strategies? *Frontline Learning Research*, *8*(3), 26–39.
https://doi.org/10.14786/flr.v8i3.489
Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*, *12*(3), 229–238.
https://doi.org/10.1023/A:1023254226592
Conway, A., Kane, M., Bunting, M., Hambrick, D., Wilhelm, O., & Engle, R. (2005, October 1). Working memory span tasks: A methodological review and user's guide [journal article]. *Psychonomic Bulletin & Review*, *12*(5), 769–786.
https://doi.org/10.3758/BF03196772
Donche, V., & Van Petegem, P. (2008). The validity and reliability of the short inventory of learning patterns. In E. Cools, H. van den Broeck, & T. Redmond (Eds.), *Style and cultural differences: How can organisations, regions and countries take advantage of style differences*(pp. 49–59). Leuven University Press.Duchowski, A. (2007). *Eye tracking methodology: Theory and practice*. Springer.
https://doi.org/10.1007/978-1-84628-609-4
Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*(2), 211–245.
https://doi.org/10.1037/0033-295X.102.2.211
Fowler, F. (2014). *Survey research methods - (5th edition* ed.). Sage publications.Fowler, F., & Cosenza, C. (2008). Writing effective questions. In E. de Leeuw, J. Hox, & D. A. Dillman (Eds.), *International Handbook of Survey Methodology*. Taylor & Francis Group 136-160 .Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, *73*(2), 349–360.
https://doi.org/10.1093/poq/nfp031
Galesic, M., Tourangeau, R., Couper, M. P., & Conrad, F. G. (2008). Eye-tracking data: New insights on response order effects and other cognitive shortcuts in survey responding. *Public Opinion Quarterly*, *72*(5), 892–913.
https://doi.org/10.1093/poq/nfn059

Gathercole, S., & Alloway, T. (2013). *De invloed van het werkgeheugen op het leren: Handelingsgerichte adviezen voor het basisonderwijs*. SWP.Graesser, A., Cai, Z., Louwerse, M., & Daniel, F. (2006). Question Understanding Aid (QUAID) a web facility that tests question comprehensibility. *Public Opinion Quarterly*, *70*(1), 3–22. https://doi.org/10.1093/poq/nfj012

Helgeson, J. G., & Ursic, M. L. (1994). The role of affective and cognitive decision-making processes during questionnaire completion. *Psychology & Marketing*, *11*(5), 493–510. https://doi.org/10.1002/mar.4220110506

Herzog, A. R., & Bachman, J. G. (1981). Effects of questionnaire length on response quality. *The Public Opinion Quarterly*, *45*(4), 549–559. www.jstor.org/stable/274890310.1086/268687

Höhne, J. K. (2019). Eye-tracking methodology: Exploring the processing of question formats in web surveys. *International Journal of Social Research Methodology*, *22*(2), 199–206. https://doi.org/10.1080/13645579.2018.1515533

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.Hyönä, J., & Lorch, R. F. (2004). Effects of topic headings on text processing: Evidence from adult readers' eye fixation patterns. *Learning and Instruction*, *14*(2), 131–152. https://doi.org/10.1016/j.learninstruc.2004.01.001

Hyönä, J., Lorch, R., & Rinck, M. (2003). Eye movement measures to study global text processing. In R. Hyönä (Ed.), *The Mind's Eye: Cognitive and applied aspects of eye movement research* (pp. 313–334). Elsevier Science. https://doi.org/10.1016/B978-044451020-4/50018-9

James, J. M., & Bolstein, R. (1990). The effect of monetary incentives and follow-up mailings on the response rate and response quality in mail surveys. *Public Opinion Quarterly*, *54*(3), 346–361. https://doi.org/10.1086/269211

Just, M., & Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329. https://doi.org/10.1037/0033-295X.87.4.329

Just, M., & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122. https://doi.org/10.1037/0033-295X.99.1.122

Karabenick, S., Woolley, M., Friedel, J., Ammon, B., Blazevski, J., Bonney, C., De Groot,E., Gilbert, M., Musu, L., Kempler, T., & Kelly, K. L. (2007). Cognitive processing of self-report items in educational research: Do they think what we mean? *Educational Psychologist*, *42*(3), 139–151. https://doi.org/10.1080/00461520701416231

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, *2*(1), 15–47. https://doi.org/10.1016/0010-0277(72)90028-5

Krosnick, J. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, *5*(3), 213–236. https://doi.org/10.1002/acp.2350050305

Krosnick, J., & Alwin, D. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly*, *51*(2), 201–219.

https://doi.org/10.1086/269029
Lenzner, T., Kaczmirek, L., & Galesic, M. (2011). 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.
https://doi.org/10.1093/ijpor/edq053
Lenzner, T., Kaczmirek, L., & Galesic, M. (2014). Left feels right: A usability study on the position of answer boxes in web surveys Social Science Computer Review . *32*(6), 743–764.
https://doi.org/10.1177/0894439313517532
Lenzner, T., Kaczmirek, L., & Lenzner, A. (2010). Cognitive burden of survey questions and response times: A psycholinguistic experiment. *Applied Cognitive Psychology*, *24*(7), 1003–1020.
https://doi.org/10.1002/acp.1602
Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, 1171.
https://doi.org/10.3389/fpsyg.2015.01171
MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on just and carpenter (1992) and waters and caplan (1996). *Psychological Review*, *109*(1), 35–54.
https://doi.org/10.1037/0033-295x.109.1.35
Meade, A., & Craig, S. (2012). Identifying careless responses in survey data. *Psychological Methods*, *17*(3), 437.
https://doi.org/10.1037/a0028085
Menold, N., Kaczmirek, L., Lenzner, T., & Neusar, A. (2014). How do respondents attend to verbal labels in rating scales? *Field Methods*, *26*(1), 21–39.
https://doi.org/10.1177/1525822X13508270
Miller, K. (2014). Introduction. In K. Miller, S. Willson, V. Chepp, & J. L. Padilla (Eds.), *Cognitive interviewing methodology*. John Wiley & Sons, Inc 1-5 .Neuert, C. (2016). *Eye tracking in questionnaire pretesting*.Neuert, C. E. (2021). The effect of question positioning on data quality in web surveys. *Sociological Methods & Research*, 004912412098620.
https://doi.org/10.1177/0049124120986207
Neuert, C., & Lenzner, T. (2016). Incorporating eye tracking into cognitive interviewing to pretest survey questions. *International Journal of Social Research Methodology*, *19*(5), 501–519.
https://doi.org/10.1080/13645579.2015.1049448
Olsen, A. (2012). The Tobii I-VT fixation filter.Olsson, P. (2007). Real-time and offline filters for eye tracking.Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422.
https://doi.org/10.1037/0033-2909.124.3.372
Schuman, H., & Presser, S. (1996). *Questions and answers in attitude surveys: Experiments on question form, wording, and context*. SAGE Publications.
https://books.google.be/books?id=Je640UKqNaYC
Schwarz, N. (1990). Assessing frequency reports of mundane behaviors: Contributions of cognitive psychology to questionnaire construction Hendrick, C, Clark, M.S. In *Research methods in personality and social psychology*. Sage Publications 98-119 .Schwarz, N. (2007). Cognitive aspects of survey methodology. *Applied Cognitive Psychology*, *21*(2), 277–287.
https://doi.org/10.1002/acp.1340

Sudman, S., Bradburn, N., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. Jossey-Bass.Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Rational thinking and cognitive sophistication: Development, cognitive abilities, and thinking dispositions. *Developmental Psychology*, *50*(4), 1037–1048. https://doi.org/10.1037/a0034910

Tourangeau, R. (1984). Cognitive sciences and survey methods. In T. Jabine, M. Straf, J. Tanur, & R. Tourangeau (Eds.), *Cognitive aspects of survey methodology: Building a bridge between disciplines* (pp. 73–100). National Academy Press.Tourangeau, R. (2003). Cognitive aspects of survey measurement and mismeasurement. *International Journal of Public Opinion Research*, *15*(1), 3–7. https://doi.org/10.1093/ijpor/15.1.3

Tourangeau, R., & Bradburn, N. (2010). The Psychology of Survey Response. In P.Marsden & J. Wright Eds. *Handbook of Survey Research. Second Edition* (2nd). EmeraldGroup Publishing Limited 315-346 .Tourangeau, R., & Hanover, L. (2018). The survey response process from a cognitive viewpoint. *Quality Assurance in Education*, *26*(2), 169–181. https://doi.org/10.1108/QAE-06-2017-0034

Unsworth, N., Heitz, R., Schrock, J., & Engle, R. (2005, August 1). An automated version of the operation span task [journal article]. *Behavior Research Methods*, *37*(3), 498–505. https://doi.org/10.3758/BF03192720

Vannette, D., & Krosnick, J. (2014). A comparison of survey satisficing and mindlessness. *The Wiley Blackwell Handbook of Mindfulness*, *1*, 312 https://doi.org/10.1002/9781118294895.ch17

.Vermunt, J. D. H. M., & Donche, V. (2017). A learning patterns perspective on student learning in higher education: State of the art and moving forward. *Educational Psychology Review*, *29*(2), 269–299. https://doi.org/10.1007/s10648-017-9414-6

Willis, G. (2004). *Cognitive interviewing: A tool for improving questionnaire design*. Sage Publications.Willis, G., & Miller, K. (2011). Cross-cultural cognitive interviewing: Seeking comparability and enhancing understanding. *Field Methods*, *23*(4), 331–341. https://doi.org/10.1177/1525822X11416092

Yeari, M., Oudega, M., & van den Broek, P. (2016). The effect of highlighting on processing and memory of central and peripheral text information: Evidence from eye movements. *Journal of Research in Reading*, *40*(4), 365–383. https://doi.org/10.1111/1467-9817.12072