Benefits of using the eye-tracking method for qualitative observation of students' multiple-choice physics tasks solution process

Martina Kekule*, Iva Spanova*, Jouni Viiri+

*Charles University, Faculty of Mathematics and Physics, Department of Physics Education
+University of Jyväskylä, Department of Teacher Education

Abstract: As previous studies have shown, the eye-tracking method is useful for monitoring cognitive processes. In this study, we focus on the benefits of analysing high school students' problem-solving processes when completing a multiple-choice test. Specifically, we provide a qualitative analysis of eye-movements of two *low prior knowledge* students as they solve two physics tasks. Both students scored zero on the test and chose the same options. In a paper-pencil administration of such a test, we would receive the same results from both of them. Via analysis of the eye-tracking data, we have shown that there are several similarities and differences between their solving approaches. Particularly, they both emphasized the key terms in the stem, meaning they did not provide incorrect answers due to improper stem reading. Both students showed very typical misconceptions concerning understanding Newton's First Law. However, they differ in their confidence. Eve arrived at her answer quicker, but Jane needed to reread the stem many times. We observed atypical saccadic eye movements when Eve read the stems of the tasks and much longer fixations when she suddenly switched to her answers. Jane showed "graph as a picture error" when interpreting an illustrative picture in the stem. Based on our analysis, we demonstrate that summarized data on areas of interest (AOIs) can be misleading when AOIs are too large (e.g., each test option). We also introduce several indicators appropriate for a qualitative analysis, and we present new visualisations for in-word and between-words saccades.

Keywords: eye-tracking, physics education, regressive saccades, multiple-choice test, problem-solving

Multiple-choice question examinations are a very common way of assessment, not only in physics education. Administration of multiple-choice examinations in a paper-pencil format enables researchers to observe the

option which students chose, but not the processes used in selecting their answer. In order to gain deeper insight into students' solving processes, the eye-tracking method can be used. In contrast to an interview, it allows maintaining the economic efficiency of multiple-choice tests, which can be important for inter/national exams and in everyday teaching practice. Using eye-tracking method for educational purposes has benefits beyond following students' solving strategies. For example, one project on intelligent "anticipating" textbooks (e.g., HyperMind, 2018) aims to create educational texts which will, based on students' eye movements, adjust the text according to students' needs. This exemplifies the need for interpretation of the gaze data during common students' educational activities.

Eye-tracking technology has been proven useful in educational research in general (Lai et al., 2013; Šmideková, 2018). Several studies focused directly on following students' eye movements when solving multiple-choice tasks in reading (Bax, 2013) and in science education (Kozhevnikov et al., 2007; Madsen et al., 2012; Tsai et al., 2012). In the science education field, quantitative analysis of particular eye-tracking metrics in defined areas of interest (AOI) has been used. This approach provides efficient summarized results and is appropriate where two or more participant groups are compared, such as in the expert-novice paradigm. However, a qualitative analysis of such data can provide deeper insight and more information about the processes which are lost by data summarization.

One strand of eye-tracking educational research has focused on the comparison between high prior knowledge (HPK) students and low prior knowledge (LPK) students (Lindner et al., 2014). We believe that focusing on differences among LPK students can provide an even deeper insight into students' physics concepts.

When it comes to solving multiple-choice tests, the process includes the following students' activities:

– reading, especially reading the stem, which is usually in the written form;

– interpretation of options, which can be presented in various representations;

– cognitive strategy aiming to choose the correct option, i.e. the decision-making process (Lindner et al., 2014).

Previous studies concerning consumer decision behaviour, line-up identification or face preference decisions showed a "shift of attention towards the chosen option as a stable phenomenon" and so proven the eye-tracking method as an appropriate research method in the fields (Lindner et al., 2014). In the educational context, this "gaze bias effect" is influenced by students' domain knowledge. High prior knowledge students will focus on information more relevant for correctly solving a task. The "amount of attention paid to correct answer options is a reliable indicator for the MC (multiple choice) test performance of HPK" (Lindner et al., 2014). HPK students also differ from LPK students, since they pay "almost as much attention to attractive options as to chosen options, whereas LPK students paid much more attention to chosen than to attractive options" (Lindner et al., 2014).

Reading is an essential part of the completion of multiple-choice tests. Many studies have shown that the eye-tracking method is useful in monitoring cognitive processes concerning reading (Dimigen et al., 2011). Hereafter, we will focus only on "natural reading" in contrast to reading with gaze-contingent display (Kliegl et al., 2006, p. 7). Using eye-tracking, reading can be distinguished from not reading and from skimming. Furthermore, differences between experts and novices in cognitive processes when solving various problems can also be observed (Gegenfurtner et al., 2011). This paradigm supposes that strategies used by experts are the more effective ones.

In our previous study, we focused on students' cognitive processes while solving multiple-choice test from mechanics. Based on summarized data in defined areas of interest (stem and each option) we observed differences in approaches between correctly and incorrectly answering students (Kekule & Viiri, 2018). This is a very common result in the expert-novice paradigm (e.g. Tsai et al., 2012; Madsen et al., 2012). However, we observed differences in attention allocation between incorrectly answering students as well. In this paper, we will provide a detailed qualitative analysis of eye movements of two incorrectly answering students who achieved the same test score and chose the same options as their answers.

Our study has the following specific aims:

1) In the theoretical part, we will provide an overview of indicators appropriate for qualitative analysis of students' solving multiple-choice

tasks. As the stem is usually administered in a written form, this includes students' reading. In particular, we will present indicators based on fixations and based on saccadic eye movements. Options can be presented in different modes of representations, which will be considered.

2) In the research part, the application of these indicators to real gained eye-tracking data and providing qualitative analysis of students' solving approaches. Discussion of the limits of the application. Presentation of information enrichment when eye-tracking method instead of only paper-pencil test administration is used. Particularly, we will analyze data of two *low prior knowledge* students who solved multiple-choice test tasks and chose the same and incorrect options.

3) Presentation of appropriate and also new ways of visualisation of saccadic movements. As Blascheck et al. (2014) states, visualisation techniques allow analysing eye-tracking data in an explorative and qualitative way. Therefore, the visualization aspect of data in a qualitative approach is useful.

1   Indicators of students' cognitive processes while solving a multiple-choice task based on the eye-tracking data

Eye movements are influenced by different textual and typographical features and changes in the observed eye movements can indicate different processing of the material. Moreover, solving a multiple-choice task is a decision-making process (Lindner et al., 2014), which includes several parts (See Table 1). Firstly, it includes the general strategy of solving the task and choosing the option. Students can read the whole stem and then go from one option to another, or they can skip to the options before they complete the reading of the stem (See Figure 1). Secondly, it involves the interpretation of the task stem and of each option. A task stem is presented prevalently in the written form, so its interpretation is done through reading. On science tests, multiple choice test options are commonly displayed using various representations, such as graphs, diagrams, or equations. From the eye-tracking point of view, we can characterize the type of task as a 2D static stimuli with passive content which can be both point- or area-of-interest-(AOI)-based (Blascheck et al., 2014).

Table 1

Parts of solving a multiple-choice task process.

Part	Description of the cognitive activity	Type of data (Blascheck et al., 2014)	Indicators based on
General strategy	task/stem interpretation	AOI-based	transitions among a stem and options
	arriving at the answer		
	decision-making process about option choice		
Interpretation of			
stem	Reading	point-based	fixation and saccades
options	Reading	point-based	fixation and saccades
	interpretation of non-textual representations	AOI-based	heat maps, fixations, saccades

Analysis of the students' dealing with the task stem means analysis of students' reading. In basic terms, our vision consists of two types of events: fixations and saccades. During fixations we extract information from the material. They are typically longer than saccades which last approx. several milliseconds and serve as a transition to another fixation. Both fixations and saccades can serve as indicators for (not only qualitative) analysis of students' reading. For example, Rayner (1998, p. 376) states that "as the text becomes conceptually more difficult, fixation duration increases, saccade length decreases and the frequency of regressions increases".

A general strategy of dealing with a multiple-choice task can be visible through transitions among the stem and options. In this case, for the analysis it is useful to mark the stem and options as areas of interest (AOIs) and relate appropriate indicators to these AOIs. It is more common to work with these AOIs in the quantitative research approach; however, some outputs can serve as appropriate information for a qualitative description as well.
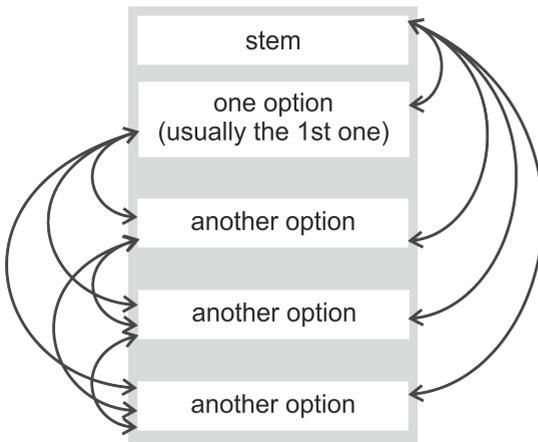
Figure 1. An example of general strategy of solving a multiple-choice task. The arrows mark the student's gaze transitions between the stem and the options.

Students' process of interpretation of each option depends on their representation. For options in the written form, indicators relating to the students' reading can be applied. For options using other representations, the areas which students paid the most attention to can be analysed and visualised by heat maps, cluster analysis, or again within AOIs approach.

Next, in line with Aim 1, we will describe selected indicators that can serve as a basis for qualitative analysis of students' eye movements when they solve a multiple-choice task.

1.1   Reading – indicators based on fixations

longer single fixation duration: l-fsd. As stated above, fixation duration increases when a text is conceptually more difficult (Holmqvist et al., 2011, p. 378). The first fixation on a word and fixations below but not above 140 ms seems to be connected with lexical activation in contrast to other reading processes. Also, earlier fixations on a picture are shorter than the later ones.

We can identify fixations longer or shorter than usual both by qualitative visualisation and by quantitative computation. Qualitatively, increasing

fixation duration can be identified using a histogram of eye fixations under the text line (Kunze, 2013a). This can reveal challenging words, for example. Quantitatively, we can set up a threshold and identify fixations above the threshold. Average fixation duration is a logical option for the threshold. During natural reading, we can expect very similar durations of each fixation and, only in some cases, increasing and decreasing values. Usually, the average value will be slightly above the most common value and will separate longer fixations. Adequacy of the threshold can be confirmed using a bar graph of all fixation durations.

It is important to take into account, that "different people have different average fixation durations" (Holmqvist et al., 2011). When we qualitatively look at students' reading of a specialized text (e. g. from physics field), it would be appropriate to include reading of common fiction into the research design as well and compare fixations duration among the two different types of text (Kekule & Krejčí, 2019). For example, Rayner and Pollatsek (1989) found that average fixations durations range from 200 ms for light fiction up to 260 ms for physics and biology texts (Holmqvist et al., 2011, p. 382). These values will vary for non-English languages.

1.2 Reading– indicators based on saccades

When reading a left-to-right language text, we proceed onwards by saccades from left to right. These saccades suggest fluent reading or possibly scanning of the text. During reading, we can identify shorter saccades than during scanning (Holmqvist et al., 2011, p. 267). When we smoothly comprehend the read text, sometimes called the "default mode of reading", our eyes will fluently move forward, with relatively few regressions (Reichle et al., 2009, p. 9). When the text is more difficult, we make more backwards saccades/ regressions, which can indicate cognitively challenging processing. These reverse saccades are usually preceded by a longer than average fixation (Tatler & Vincent, 2008 in Holmqvist et al., 2011). According to Rayner (1998) we can distinguish four types of backwards saccades. Two of them are well defined and easily identified:

few letters back within a word: s-in. Sometimes called *in-word regression* (Holmqvist et al., 2011), this is a small size regression which serves to reprocess a word that probably has not been properly analysed during the first reading (Rayner, 1998, p. 376). It can also indicate recognition of the

word as a key or important word in the context. For example, a student provide s-in on the "rozpadá" (English: decay) word in physics text focused on nuclear transformation theme (Kekule & Krejčí, 2019).

few words back: s-btwn. Also called *between-word regression* (Holmqvist et al., 2011), this means returning back within a sentence and indicates the failure to understand the content (Rayner, 1998, p. 376).

For the other two types, various names are used in eye-tracking literature. We will describe the two types which differ distinctively from one another and we will discuss their possible connection with more detailed definitions in other references.

precise return: s1. Eye's returns to a precise fixation point are sometimes called 'return sweeps' (Bax, 2013, p.7). These saccades are typical for higher-proficiency readers, who are able to accurately find the position in the text that caused them difficulties and that they need to focus on. This type of saccade can also indicate a student's ability to interconnect information from the text (Kekule & Krejčí, 2019). For example, a student reading about alpha decay and changes in atomic and nucleon number (see Fig. 2 left, black rectangle) returned to the information about the numbers for beta decay in the text (see Fig. 2 left, gray rectangle). Moreover, students provide this type of saccade when, after reading a text, they are searching for an answer to a question (Kekule & Krejčí, 2019; See Fig. 3). Holmqvist et al. (2011) cite definition of the ´sweep´ by Aaltonen et al. (1998) as "a sequence of saccades that move in the same direction."

returning back: s2. In contrast, low-proficiency readers do not return precisely to the point, but they backtrack through the text until they find what they were searching for (Bax, 2013, p. 7). We can observe this type of not precisely focused regressive movements when students try to find an answer to a posed question in the read text, but they do not remember where the information was.
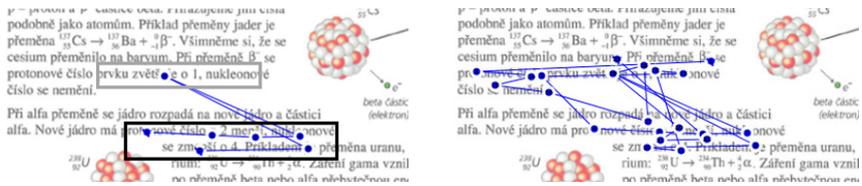
Figure 2. Precise return s1. A student reading about alpha decay and changes in atomic and nucleon number (left, black rectangle) returned to the information about the numbers for beta decay in the text (left, gray rectangle). He repeated his action several times (right). Illustrative figure based on Kekule & Krejčí (2019).

Note: Saccades in the opposite direction will be identified when a reader is going to read the new line of a text from the previous line end. This type of saccade, the so-called line break (Kunze, 2013b) or a type of look-ahead scan-path event (Holmqvist et al., 2011), will be omitted from our analysis.
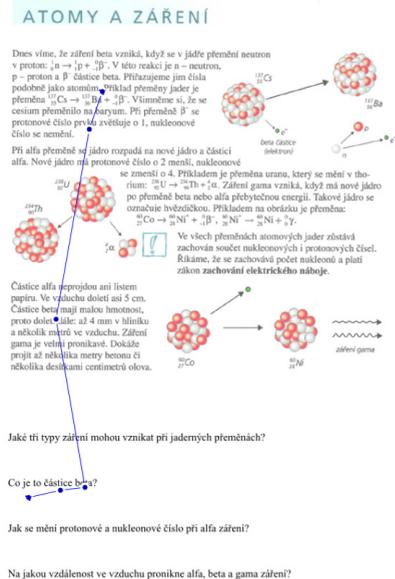
Figure 3. Precise return s1 saccade after reading a question: What is a beta particle? A student after two fixations arrived at the part of the text about the beta decay. Illustrative figure based on Kekule & Krejčí (2019).

1.3   Indicators for areas of interest (AOIs)

summarized data: total dwell time on AOI – tdt.
summarized data: fixations count on AOI – fc.
summarized data: average fixation duration on AOI – fad.

These three indicators, which are based on a larger number of fixations, can inform us either about attraction of an AOI or performance under an AOI. The number of fixations prior to first entering an AOI is a good indicator of its noticeability; the number of fixations on an AOI is an appropriate indicator of interest in the AOI. Average fixation duration is a measure of cognitive processing (Bojko, 2013, p. 125). These indicators can be used without defined AOIs as well. In that case, fixation density can be explored, which shows the number of fixations on cm2 (Holmqvist et al., 2011, p. 403). Total attention paid to an AOI can be described by total dwell time on an AOI, i.e., the sum of all dwell times spent on an AOI (Holmqvist et al., 2011, p. 389). Definitions of dwell time can vary within a literature, as do its names, such as visit/observation/gaze time. It is often defined as the sum of all fixation durations during a visit on an AOI (Holmqvist et al., 2011, p. 389). Sometimes it is counted based on raw data; in that case saccadic durations are included as well. However, fixations make up most of the gaze time, so both approaches are acceptable. In this study, we adopt the first approach and define dwell time as the sum of all fixation durations during an AOI visit.

Besides these summarized data, we can take into account also the time dimension and the sequences of AOIs – i.e., the order, in which a person visited each AOI and the transitions among them. For qualitative analysis of eye movements, we can use the following parameters as indicators:

Occurrence and order of an AOI within the sequence;
Significant frequency of transitions between two AOIs.

We can produce a simple sequence, which can serve as a basis for transitions analysis. Firstly, we list the order in which AOIs were visited (see Fig. 4a)). Holmqvist et al. (2011, p. 193) label this sequence as a *compressed string*. Secondly, we can add information about the number of fixations on the AOI during the visit (see Fig. 4b); termed as *string* (Holmqvist et al., 2011, p. 193). Thirdly, we can add information about fixations duration on the AOI during the visit and create a so called AOI *sequence chart* (see Fig. 4 c); Holmqvist et al., 2011, p. 193–200).

a)A B D A C B C D C D A
b) 5A 3B 3D 10A 2C 2B 3C 5D 2C 3D 5A or AAAAABBBDDDAAAAAAAAAACCBBCCC.....
c) AOI sequence chart. Horizontal axis shows viewing time and vertical axis shows AOIs.
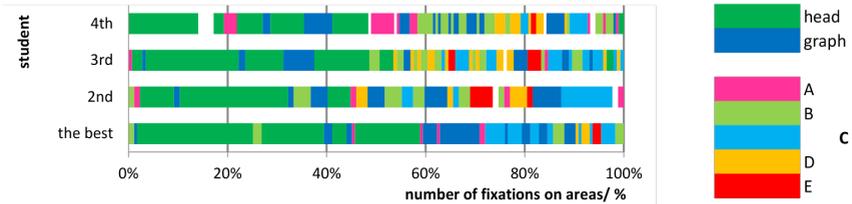
Figure 4. Ways of AOIs sequences representations. AOI sequence chart created with BeGaze software (Viiri et al., 2017).

We can summarize string data into absolute or relative number of transitions between two AOIs as shown in Figure 5. Moreover, Blascheck et al. (2013) presents a "circular heat map transition diagram" as an appropriate tool for transitions visualizations of a participant. The AOIs are segments on a circle, where their size corresponds to the dwell time on each AOI, and transitions are visualized by colored areas among the segments. These diagrams are appropriate for example for comparison of participants' strategies. Figure 6 shows an idea of a similar diagram produced with Eyetrace software.
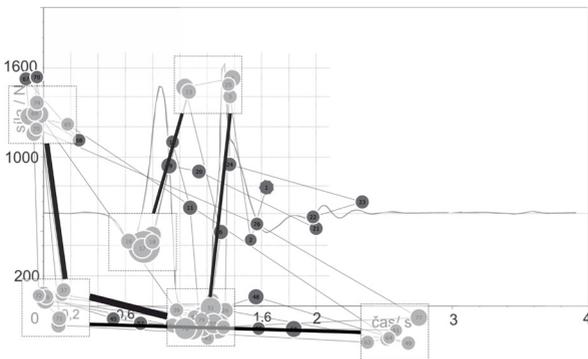
Figure 5. Transitions among AOIs. Absolute and relative count. Ogama software (Voßkühler et al., 2008).
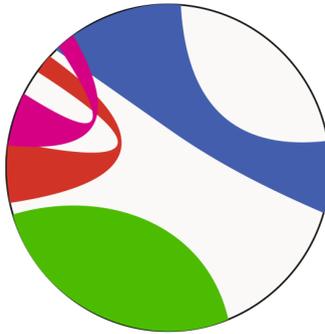
Figure 6. Transitions among AOIs. Idea of the circle diagram which can be produced by the Eyetrace software (inspired by Kübler et al., 2015).

We can distinguish between one-way and two-way transitions (Holmqvist et al., 2011, p. 191). The latter one means going back and forth between two AOIs and is a sign of "actual intergration" (Holmqvist et al., 2011, p. 191). Transition rate, or how many transitions are made per a second, could be another measure. Rate threshold can indicate whether a participant is able to perform a specific task, such as reading lyrics and music score when singing (Holmqvist et al., 2011, p. 424). In such case, the task demands a minimal transition speed between the lyrics and music score. However, when a participant can switch between two areas of interest freely, according to Holmqvist et al. (2011, p. 424)) "there is no consensus that a higher transition rate … implies better integration between the two information sources."

Sometimes students go to an empty space when they are thinking about a possible answer. Where they return after visiting these "out" areas can serve as information about ongoing cognitive processes as well.

Holmqvist et al. (2011) distinguishes so-called:

   look-backs. These are saccades which return to AOIs already looked at. Sometimes they are called refixations as well. Here a question about what it means "to return to the content of AOI" arises. When the content of the visited AOI is no more in the working memory, we should consider the look-back rather as a "first visit". For example, we can then limit the look-backs to a 10 s window (Holmqvist et al., 2011). If we compare this term with definition in part 1.2, we can include s1 and s2 indicators into the look-backs category. As

we show in Fig. 3, after reading larger text for more than 10 seconds, some students are able to remember key terms which they read about and look them up quite easily.

2   The test session

2.1   Testing procedure and participants

Participation in the research was voluntary, anonymous and with respect to research ethic. Students shortly got to know the eye-tracking device. After the device was calibrated, they started solving the test tasks. The eye-tracking session lasted about 8 minutes. During the whole session the participants were asked to think aloud when it was needed; the sessions were videotaped. Afterwards, the participants were asked to provide an oral retrospective report and to complete a learning style preference questionnaire (see below).

More than 40 high school students and newly enrolled first-year university students took part in the study; data from 35 participants were valid and included in further analysis. In this study, we focus on two students of these participants.

2.2   Test & Questionnaire

The test consisted of four tasks from mechanics, specifically three tasks from the R-FCI test (Nieminen et al., 2010) and one task originally created for the purpose of the test and following the same pattern as the three previous ones. Two tasks focused on understanding of Newton's 1st Law and two of Newton's 2nd Law. (The tasks are shown in the Appendix). Both laws are conceptually challenging for students and typical non-Newtonian conceptions are commonplace (Hestenes et al., 1992). Options were presented in either verbal or graph representation. All task stems appeared in written form.

We will provide analysis of two of the tasks. Both tasks dealt with the same context, rocket in the outer space, which allows for a better comparison. The tasks differ in their aims: the first ("rocket before" task) was focused on the understanding of *Newton's 2nd law*, whilst the other one ("rocket after") to *Newton's 1st law*. In addition, they differ in representation of offered options: options of "rocket before" are in written form, options of "rocket after" are in graph representation. Both tasks contained an additional picture, which served a mostly decorative purpose. "Rocket before" was the first task,

which students solved, then "rocket after" followed. Only the two tasks are compared due to the limited extent of the paper. These two tasks use the same context and conceptually complement each other in such a way that their analysis enables us to take a more detailed look at the students's conceptions of Newtonian mechanics.

Learning preferences were investigated by a part of the Learning Style Inventory (Dunn et al., 1989) in the Czech adaptation (Mareš & Slavik, 1990). The questionnaire consists of statements to which participants express the extent of their agreement on the 5-point Likert scale and focuses on the following dimension: audio/visual learning, tactile and kinaesthetic learning. It was included as the preference of the visual learning style could have impacted ours results as we follow students' visual attention. Both selected students' results show preference for kinaesthetic learning style. Detailed description of the test procedure and all tasks included in the test is available in a previous paper (Kekule & Viiri, 2018).

2.3   Equipment

Eye-tracker Tobii TX300 with frequency 300 Hz, which has an accuracy less than 0.5° of visual angle, was used. The infrared camera was placed under the 23-inch screen of the stimulus PC. Stimuli were presented as a pdf-document. Participant's eyes were positioned at a distance of approximately 70 cm from the centre of the screen. A five-point calibration and validation procedure was used before the start of the experiment. Eye movements were recorded by Tobii Studio 3.2 and for identification of fixations inbuilt IVT filter was used. Eye movement was classified as a saccade when eye's velocity exceeded 30 °/s. Minimum fixation duration was set to 60 ms. An extra mouse and keyboard were connected to the stimulus PC by which participants handled the PC.

3   Jane & Eve

In this section, we will provide qualitative analysis of recorded eye-movements of two girls who participated in the study described in part 2. General characteristics of the girls based on the questionnaire data and test results are provided below. Our aim is to show that while the girls tend to show similarities in the test, recordings of their eyes when they solved the tasks were different.

Jane is a 17-year-old high school student, participant no. 12. She attends a high school for gifted students in Prague. Her last grades in both mathematics and physics were average (grade 3, where 1 is best and 5 is worst). She rather agrees that she prefers science to art/humanities. She finds physics useful. As for learning preferences, she prefers neither audio nor visual approach to learning, but learns by building something, hands-on experiments, etc. She appreciates emotional engagement when she learns but does not like to move or express herself by movement.

Eve is an 18-year-old high school student, participant no. 39. She attends high school for gifted children in Prague. Her last grades in both mathematics and physics were better than average (grade 1 and 2). She agrees that she prefers science to art/humanities. She enjoys physics and finds it useful. As to learning preferences, she prefers visual learning to listening or talking about the content with someone. She rather prefers to learn by building something and by hands-on experiments. She enjoys experiential type of learning the most.

Jane needed slightly more time (4:59 min) to go through the whole test than Eve (3:41 min). Both girls chose the same options, all of which were incorrect. Jane was solving the first task almost two times (109 s) longer than Eve (56 s). Detailed data are provided in Table 2.

Table 2

Test results for both participants Jane and Eve. Tasks from left to the right respect order in the test setup.

Task	Rocket before	Astronaut	Rocket after	Woman and box
Focused on	Newton's 2nd law	Newton's 1st law	Newton's 1st law	Newton's 2nd law
Representation of the options	text	graph	graph	motion map
Jane				
Chosen option	e	b	c	e
Viewing time/s	109.28	54.57	53.87	81.12
Eve				
Chosen option	e	b	c	e
Viewing time/s	56.00	68.27	33.25	63.43

4   "Rocket before" task

The original test layout of the "rocket before" task is in Appendix I. The task required qualitatively determining a rocket's speed under certain conditions. In short, the task posits that only one constant force is acting on the rocket. Students can choose from these options: a) the speed is constant for a while and decreasing thereafter, b) increasing for a while and constant thereafter, c) continuously decreasing, d) continuously increasing and e) constant. The correct answer is d) "continuously increasing".

4.1   Reading the stem

Jane chose the last option "constant" as her answer. Gaze plot is available in Apendix III.

Fixations 4 to 64 were focused on the stem. Then she was focused on the options as well. She read the stem fluently with several increasing fixations and regressive saccades. Average fixation duration when reading the stem was 251 ms. We will use this value as the threshold to identify longer fixations. A bar graph in Appendix II showing fixations duration proves the threshold as a reasonable value.

Fixations no. 4, 7, 9, 12, 13, 14, 16, 19, 20, 21, 22, 24, 30, 50, 52, 54, 57, 63, and 64 lie above the threshold. Fig. 7 shows positions of these fixations.

Figure 7. Jane's fixations with longer than average duration. Each circle shows one event. Original layout in Czech language.

Figure 8. Eve's fixations with more than average duration. Each circle shows one event.

To understand Jane's fixations, they are shown below in the English translation of the task. Syllables with long fixations are in bold.

A **roc**ket drifts sideways in outer **spa**ce. The rocket is subject to **no out**side **for**ces. At the instant of **tim**e t0 the **roc**ket's engine is turned on and produces

a **con**stant force on the rocket. The force is in the direction of the motion. At the instant time t8 the rocket's engine is turned off.

During the time inter**val** t_0-t_8 the **speed** of the rocket is:

As we can see, Jane made longer fixation duration at the beginning of the stem. Generally, she made fixations on the key words of the task, which states conditions of the task (no outside forces are acting, motor force is constant, and that we ask about the speed). She tended to make longer fixations on the first syllable of a word. This visualisation could be similar to someone reading the text aloud and putting stress on the syllables.

Cognitively challenging part of the text can be revealed by an analysis of the saccadic movements. Particularly, we will focus on regressions. Jane read the stem quite fluently, we observed five areas where she made returns, two of them *in-word*, two of them *between-word* (in theory part named *s-in* and *s-btwn*) and one s1 or s2 return. See Fig. 9.

Figure 9. Jane's areas of identified returns. In-word (black), between-word (dashed), returning back (gray).

Figure 10. Eve's areas of identified returns. In-word (black), between-word (dashed), vertical saccades (vertical line) (comp. Fig. 12).

The *s-in* are under the words: *rockets' (engine)* and *(constant) force (in)*. For visualization of the *s-btwn* saccadic movements we used graphs, which use the vertical axis as a discrete time axis (see Fig. 11). This type of graph presents new way of visualization of the eye-tracking data, which could belong to the category of "timeline AOI visualisation" (Blascheck et al., 2014, Fig. 14). In our case, we consider each syllable as a kind of an AOI and instead of continuous time scale, we use discrete scale, where each level represents saccadic events in one of the two direction (to the right or to the left).

Jane made the first *s-btwn* in the first sentence. Particularly, under the part: *kde na ni nepůsobí žádné vnější síly*, in English translation: *where there any outer forces don't act on it*. The saccadic movements are shown in Fig. 11. The back saccadic movements start on a word *na/on*, goes to *kde/where*, then continues ahead to *ne(působí)/don't(act)*. Line 3 shows long reverse saccade along several words, etc.

The process indicates a need to re-read the important information that none other forces are involved. She made similar regressive movements under the sentence: *V čase t_8 je motor vypnut / At the instant of time t_8 the rocket's engine is turned off.* The third saccadic regression concerned the last part of the stem, where we asked about the rocket speed during time interval t_0-t_8. Here we observed several saccadic movements back in a row.
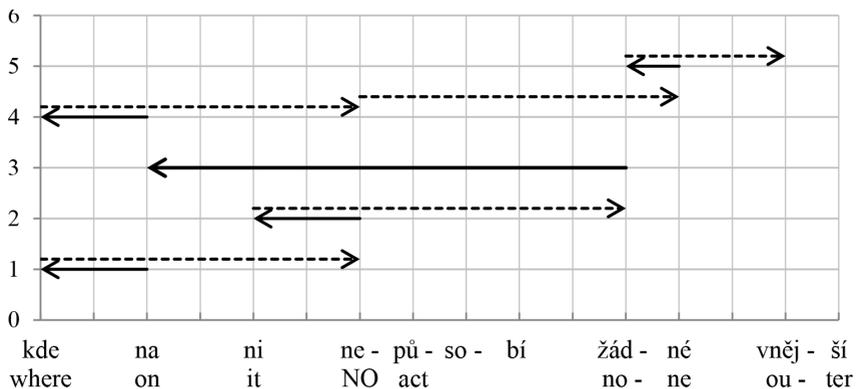
Figure 11. Jane Regressive saccadic movements under the part of the first stem sentence. Black lines – movements back, black dashed lines – movements ahead. Line 1 is a start of the whole process.

The analysis revealed that Jane paid attention to the important part of the stem and did not omit any crucial part in order to understand the task.

Eve chose the last option "constant" as her answer as well. Gaze plot is available in Appendix III. Similarly to Jane, Eve firstly carefully read the stem and then went through all the options. Fixations 2 to 78 were focused on the stem. Average fixation duration when reading the stem was 253 ms. Bar graph showing fixations duration is in Appendix II.

Altogether 27 fixations had duration longer than average. Fig. 8 shows their location on particular syllables. In contrast to Jane, Eve sometimes made longer fixations on the last syllables. She paid several times longer fixations to the key words in the last row (***Du**ring the time interval t_0-t_8 the **speed** of the ro**cket** is:*). She emphasised very similar areas as Jane did, but she did not pay longer fixation to the word *constant* (force).

In contrast to Jane and to observations in previous studies, we observed several times repeated vertical saccades under particular words (rocket, force, at time, speed, and which; see Fig. 12). This potentially points to Eve's increased cognitive processing as well. Similarly to Jane, Eve made several *s-in* and one *s-btwn* saccades (see Fig. 10). In English: s-in is under the words: *direction, speed of rocket. S-btwn* is under the words: *during the time interval.* In contrast to Jane, Eve does not show increasing attention to the statement that rocket engine is switched-off.
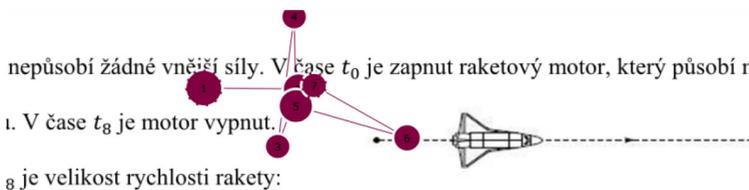
Figure 12. Eve's observed vertical saccades, which interrupted fluent reading.

4.2   Option choice & stem-choices transitions

Jane spent much more time (83 s) on the stem than on the option AOIs. Jane's total dwell time on each option AOI is shown in Fig. 13 (left). Option d) is the correct answer. She paid the least attention to option e ("constant"), which she chose as her answer and option b ("increasing for a while and constant thereafter"). As we can easily see, Eve allocated her attention to each option in a very different way (Fig. 13 (right)).

Note: We show absolute total dwell time on an AOI, though, each AOI contained text of different length. For example, option e) is the shortest, so we can expect the shortest attention on it, if someone only read the word. But option b) is the longest and as we can see, Jane also paid not so much time to it. In this case, the different length is probably not a very important parameter to consider. In case of a greater difference in length of each option, we would need to present relative values. When providing comparison between Jane and Eve, we do not need to take this into account.
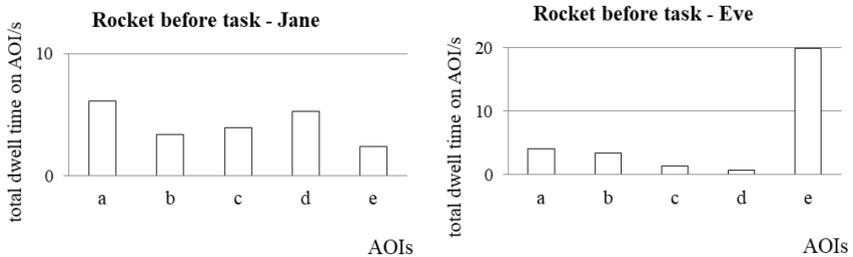
Figure 13. Total dwell time on each AOI (in s).

After reading the stem, Jane went through the options and she returned to the stem and vice versa several times. Particularly, she made 14 transitions from the stem to the options and 2 transitions from the stem to blank space. A sequence of visiting stem and options by Jane is shown in Fig. 14. First, with number of fixations, second, only visits. Stem is marked by s, option a)-e) by letters a-e.

6a 21s 1a 30s 4c 5d 2e 6s 6a 37s 1a 4b 3c 52s 2b 3c 5d 1c 1b 1a 2b 1s 3a 20s 1a 3s 1a 1s 1a 10s 1a 2b 1c 1d 1e 13s 2e 20s 1a 11s 2c 5d 1c 1a 22s 2e 1a 6s 1c 1d 1e 1a 1e 3s

a s a s c d e s a s a b c s b c d s b a b s a s a s a s a s a b c d e s e s a s c d c a s e a s c d e a e s

Figure 14. Jane's gaze sequence.

6a 15s 1c 6b 3c 2d 3e 1d 2e 1b 5a 2s 1a 4b 1d 2c 1s 2a 15e 1b 4s 5e 5s 1b 5e

a s c b c d e d e b a s a b d c s a e b s e s b e

Figure 15. Eve's gaze sequence.

Jane mostly paid few fixations to an option, and many more to the stem. We can also notice that transition stem-option a) was made very often. During the first visit of each option, she skipped option b), though, in the middle part of the solving process, she returned to it more times than to the other options. In the end, she omitted it again. If we look at Fig. 13 (left), we can see that her total dwell time (sum of all fixation durations) on AOI-b was rather short.

After the first visit of option a), she goes to stem for information: engine is switched-on and switched-off. In reality, these are two of the crucial parts for arriving at the correct answer. After first visit of option c) she continues to d) and e). After visiting e) ("constant") she goes to the picture and moves along the line – the trajectory of the rocket. It can be assumed that she shows a typical "graph as a picture" error (Beichner, 1994), when students interpret graph as a sketch of a real situation, e.g. decreasing line in a position(time) graph taken to mean that an object is falling. She then reads the stem again for a long time and gazes on option a), then for the first time on b) and then on c). After reading option c) ("continuously decreasing"), she goes to the stem and fixates the most on the sentence: *At time t_8 the engine is switched off.* This is a typical misconception connected with the Newton's 1st Law, when students expect that when there is no acting force an object will at some point come to a stop.

During sequence *20s 1a 3s 1a 1s* almost all fixations are on the time interval "t_0-t_8" and some of them are very long (3.446 s, 2.110 s, 0.756 s). It is obvious that she was aware which time interval the task is about and that she focused her attention on it.

She also compares the options with each other. For example, a sequence *b* a *b* means comparison of the beginnings of the options. Then, she read "is increasing" in option a) and again within the stem "the engine is switched off". In the second half of the whole chain of AOIs, she starts to visit option e) ("constant") more often. In the stem she tends to read the two parts about engine being switched-on and switched-off. During sequence *c d c*, she repeatedly makes horizontal saccades along each option, which suggests thinking about the options (see Fig. 16).

When comparing sequences and time allocation to each AOI, she spent the least time on option e) (her answer), which she started to be focus on in the second half of her choice process, and on option b), which she paid attention to more during the first half of the process.

b) zpočátku se zvyšuje a poté je konstantní.

c) neustále se snižuje.

d) neustále se zvyšuje.

Figure 16. Jane. The "c d c" sequence showing horizontal repetition of reading words in option d) and c).

As for Eve, after reading the stem, she went through the options and sometimes returned to the stemTotal dwell time (sum of all fixation durations) on each AOI is shown in Fig. 13 (right). She spent altogether almost the same amount of time (23 s) on the stem as on option e) ("constant"), which she also chose as her answer. The least amount of attention she paid to the correct answer – option d).

A sequence of visiting stem and options is provided in Fig. 15; first, with the number of fixations, second, only visits. Stem is marked by s, options a)-e) by letters a-e.

In contrast to Jane, Eve made only 5 transitions from the stem to options and three to blank space. Also, in contrast to Jane, she made only several fixations on an AOI and then moved on, but 15 fixations on the stem after examining option a) and 15 fixations on option e) in the second half of her decision-making process.

After reading option a) ("constant for a while and decreasing thereafter") for the first time, Eve goes to the stem in the part "rocket engine is switched on". Having read "the rocket engine is switched off", she moves to option c). This indicates the same misconception as in Jane's case.

During the first visit of option b) Eve makes a very long fixation (757 ms) near the word "constant". It was the 28th fixation after she started reading option a). Altogether she made 109 fixations on options. Then she moves along c) and d) (they do not contain the word "constant") to option e) ("constant"). Here she makes several long fixations (857 ms, 537 ms, 290 ms). Then she moves to option a). Again, on the word "constant" she makes a long fixation (657 ms), and one more when she moves to option b) (fixation duration 783 ms).

For visualisation see Fig. 17, where the words "constant" are marked within the options. It is obvious that increased values of dwell time on AOIs a and b are due to the longer fixations on the part with the word "constant". These observations highlight the fact that establishing AOIs needs careful consideration in order to interpret the summarized data correctly.

When Eve made 15 fixations on option e), she interrupted them by down gaze. Again many of these are range from 800 to 1100 ms. This is a strong indication of thinking about the option. After this, she checked the stem part about switched-on engine acting on the rocket and returned and chose option e).
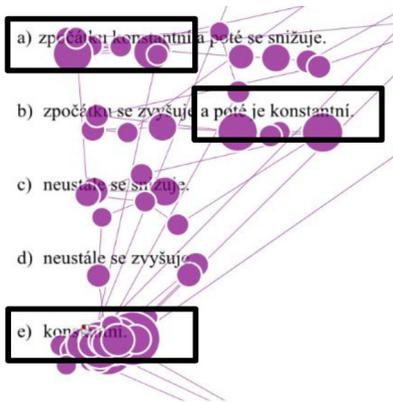
Figure 17. Gaze plot. The longer fixation the larger the circle. The areas of options, which contain the word "constant" are marked with a rectangle.

Summary

We observed similarities as well as significant differences in approaches of both girls. Both students carefully read the stem first and then began choosing an option, still looking back to the stem. Even though they chose the incorrect answer, when reading the stem, they both showed emphasis on the key words and statements crucial for arriving to the correct answer. Eve spent much less time by reading the stem and options in contrast to Jane. Eve also tended to fixate longer on the ends and not the beginnings of words. The word "constant" resonated with Eve very much and she paid the least amount of attention to the correct answer. Her first very long fixation (757 ms) happened roughly in the first third of the whole option choosing process. She had probably been thinking about the concept of "constant" as her answer since the beginning. Jane, in contrast, paid attention to option e) at the end of her selection process. She spent much longer reading the stem. She did not seem to be confident in her answer at all. Jane also provided comparison among the options from the grammatical point of view. For example, she compared beginnings of options a) and b) and c) and d), which are grammatically the same. In addition, both girls tended to connect decreasing speed with switched-off engine. This is typical for non-Newtonian understanding of Newton's 1st law. However, Newton's 1st law was the basis for the following task, "Rocket after".

5　"Rocket after" task

"Rocket after" was the third task in the test and was focused on understanding of *Newton's 1st Law*. Original test layout of the task is available in Appendix I. The task demanded qualitatively determining a rocket's speed under certain conditions, namely after a constant force stopped acting on the rocket. Students are supposed to choose from options in graph representation. The content of the graph options was the same and in the same order as in the "rocket before" task: a) the speed is constant for a while and decreasing thereafter, b) increasing for a while and constant thereafter, c) continuously decreasing, d) continuously increasing and e) constant. The correct answer is option e) "constant".

Tasks　"rocket before" and "rocket after" were very similar but had several differences. Options in the "rocket before" task were in a textual representation; options in the "rocket after" were in a graph representation. Stems of both tasks firstly describe the context of the task: A rocket is in the outer space. At a certain point in time, the rocket engine is switched on and later switched off. "Rocket before" task is focused on what will happen *before* the engine is switched off, "rocket after" task is focused on what will happen *after* the engine is switched off. Because of the similarities we could estimate students' solution of the "rocket after" task based on students' eye movements during the "rocket before" task.

As we have shown above, after reading option c) in the previous task ("steadily decreasing") both girls went to the stem to the following information: the engine is switched off. So when we ask what will happen to the rocket's speed when the engine is switched off, we could expect, that they will be attracted to option c) ("steadily decreasing"). As predicted, both students selected this answer.

5.1　Reading the stem

The stem of the task has the first and the second row the same as in the previous task concerning the rocket. The third row consists of the question itself.

Again, Jane starts by reading the stem carefully; fixations no. 3 to 9 are paid to the stem, fixations no. 10 to 12 are on option a). Then she continues by reading the stem up until fixation no. 163 with 10 fixations out of text or

picture of the task. As mentioned above, the third row of the stem was new in contrast to the previous rocket task. As we can see from the gaze plot in Fig. 18, Jane needed to go through the test question (the third row) several times and 108 fixations (more than 70% of stem fixations) were focused on it. Later, she did not return to the stem so many times as in the case of the previous task, and when we compare gaze plots of both tasks in Appendix III, we can see, that the first and the second row of the stem was now familiar to her.

Figure. 18. Jane's gaze plot of the stem. The longer the fixation, the larger the circle.

During this part, she gazed down three times: fixations 38–41 and 94–98, when she read that we ask about the rocket speed; and fixations 119–124, when she read that we ask about the speed after a particular time. After the first down-gaze she returned to the same position, after the next two gazes she returned to the beginning of the third row. This suggests thinking about the possible answer and not relying only on the choice from the options.

Average fixation duration for this reading part is 250 ms. Almost 60 fixations are above average and most of them focus on the last sentence. All this indicates increased cognitive load.

Jane made regressive saccades several times especially when reading the last row (see gaze plot in Figure 18).

As for Eve, fixations no. 3 to 50 were focused on the stem. She paid only one fixation (no. 45) to option a). The average fixation duration was 235 ms. Altogether 20 fixations were above average. Fig. 19 shows their locations. Again, we observe her tendency to make longer fixations on the last syllable. And again, she made several vertical saccades (Fig. 20.) also visible in Fig. 12. The right picture in Fig. 20 suggests that Eve made these eye movements in order to process the read text (to remember, to think about it, etc.).

Raketa letí ve vesmíru, kde na ni nepůsobí žádné vnější síly. V čase t_0 je zapnut raketový motor, který působí na raketu konstantní silou ve směru pohybu. V čase t_8 je motor vypnut.

Který z následujících grafů nejlépe popisuje závislost velikosti rychlosti rakety na čase po okamžiku t_8?

Figure 19. Eve's fixations with more than average duration. Each circle shows one event.
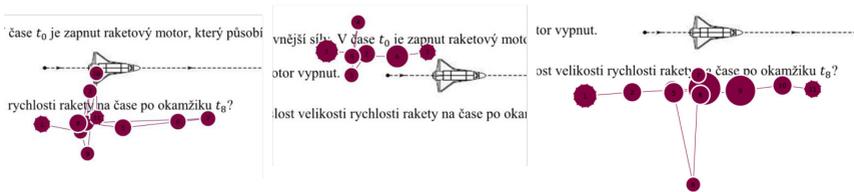
Figure 20. Eve's observed vertical saccades which interrupted fluent reading.

5.2 Option choice & stem-choices transitions

Total dwell time on each option AOI is shown in Fig. 21. Jane (left) spent much more time, altogether 39 s, on the stem than on the options. Eve (right) spent the most fixation time on option e) and option c), which she chose as her answer. On the stem she spent altogether 21 s.
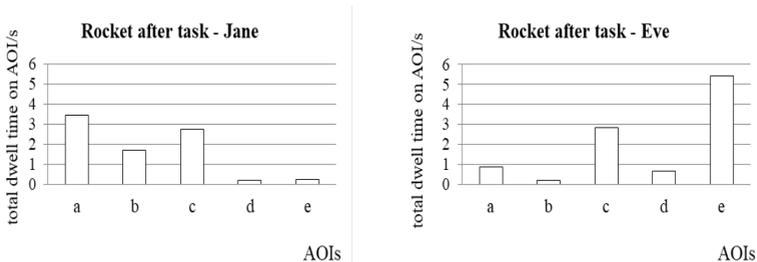
Figure 21. Total dwell time on each AOI (in s) for "Rocket after" task.

The sequence of Jane's visiting stem and options is shown in Figure 22. First, with number of fixations, second, only visits. Stem is marked by s, option a) –e) by letters a–e. She made two transitions from stem to the options.

6a 3b 4s 2a 1b 1a 1b 1a 1b 2c 1d 1e 5c 3s 2a 1b 4a 8out-to the right 5c

a b s a b a b a b c d e c s a b a out c

Figure 22. Jane's gaze sequence.

15e 1b 17s 8e 1b 3s 2d 7s 4c 1b 2a 5c

e b s e b s d s c b a e

Figure 23. Eve's gaze sequence.

Jane paid just one fixation to option d) and one to option e). After visiting options a), b) and c) she perhaps noticed, that the options are in the same order as the written ones in the previous rocket task. This simple type of graph is easily understandable in contrast to the written text. The option that she finally selected as her answer was visited more frequently. She made eight fixations to the right, out of the camera and then selected the final answer, which indicates real thinking about the task. Before this, she fixated option a) ("constant and decreasing") four times, which suggests comparing these two options.

In the beginning Jane made several transitions between options a) and b). In the stem, she fixated only on the sentence that says that engine is switched off. She was obviously aware of the key information. After that she went to the picture of the rocket and its trajectory (see Fig. 24). Then she, for the last time, returned to option a) and made the last sequence: 2a 1b 4a 8out-to-the-right 5c. All the "a" and "b" fixations are on the horizontal line of the graph and are longer; one especially long (687 ms) fixates on the horizontal line in graph a). As in the previous task, this can indicate, that she misinterprets line trajectory and horizontal (constant) line in a graph. In this case, it is also obvious that she spent more time on the a) option because of the horizontal line.

Raketa letí ve vesmíru, kde na ni nepůsobí žádné vnější síly. V čase t_0 je zapnut raketový motor, který působí na raketu konstantní silou ve směru pohybu. V čase t_8 je motor vypnut.

Který z následujících grafů nejlépe popisuje závislost velikosti rychlosti rakety na čase po okamžiku t_8?
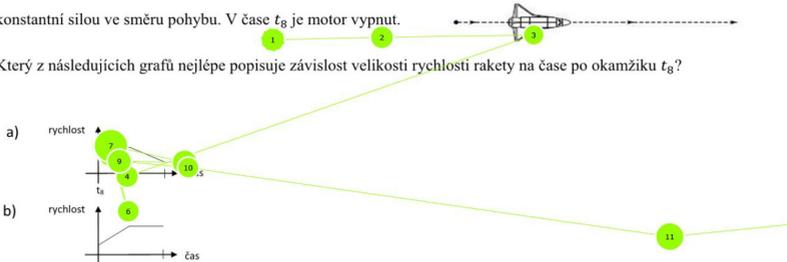
Figure 24. Gaze plot of sequence 2a 1b 4a, where Jane after fixation on the rocket went to option a) and b) and was interested in horizontal lines of the graphs.

Eve's sequence of visits to stem and options is available in Figure 23 – first, with number of fixations, second, only AOI visits. Stem is marked by s, option a)-e) by letters a-e. Similarly to Jane, she made only three transitions from the stem to the options.

As the graph in Fig. 21 shows, Eve was not much interested in option b). She visited it just once (Fig. 23). The previous fixations visible in the graph were on the edge of the area and served as stopovers from the stem to option e). She started with her favourite option e) and though the graph is very understandable, she fixated on it many times. She considered it a possible option. She returned to the third row of the stem (the question itself) and after reading it, she again clearly skipped to option e) and then went back to the stem. Specifically, she returned to the last part of the question where we asked about speed *after* the time t_8, i.e. after the engine is switched off. This seems a crucial piece of information for her and she leaves option e) and starts to be interested in the other options (see Fig. 25). The only two fixations on option a) are crucial in her decision-making process. One is the longest (490 ms) in the area just before the decreasing line of the graph. Then she moves to option c) and after several fixations she chooses it as her answer.
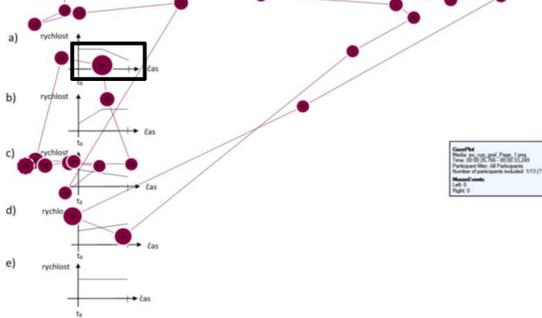
Figure 25. Gaze plots, after Eve abandons the idea about e) as the possible answer. The crucial fixation is in the rectangle (490 ms).

Summary

As in the previous task, we observed both similarities and significant differences in approaches of the students. Firstly, the task was very similar to the previous one, both girls recognized this and during the stem part, they focused predominantly on the third row, the question itself, which represented the new content. Both students seem to have been influenced by the previous task when choosing their answers. Jane made only one fixation on the two last options. At the beginning, Eve spent a lot of time on option e), which she chose as her answer to the previous task. After she noticed that we ask about slightly different information, she left the option completely.

Eve again showed longer fixations sometimes on last parts of words and vertical saccades. As shown in Fig. 20, it seems to be a sign of her further processing of what was read (remembering, thinking about, etc.).

Based on analysis of the previous task, it was to be expected that both girls would eventually choose option c) ("decreasing") as they did. According to total dwell time on AOIs, none of the girls paid the most attention to the option though. Jane paid the most attention to option a), Eve to option e).

Towards the end, both students were interested in the constant part of graphs, but for different reasons. Looking at the total dwell time data (Fig. 21), this would only be evident for Eve. Eve was obviously misled by the previous task and when she realized her mistake, she was not interested in the option e) any longer. Then she very quickly arrived at her decision about decreasing speed. Jane was again attracted by the rocket trajectory (in the illustrative picture) and that is why she was interested in horizontal lines in option b) and especially in a). After longer thinking out of the camera she decided for option c).

6   Discussion

In this study we focused on different solving processes during physics test of the two incorrectly answering students. We called them low prior knowledge (LPK) students in line with Lindner et al. (2014). The knowledge level was judged based on the results of the administered test, in which the two students scored zero. Even though the test comprised only four tasks, we can still consider the result an adequate estimation of their knowledge in the particular area of physics. We were interested in their basic knowledge of the 1st and 2nd Newton's Laws and both students showed typical non-Newtonian conception, which are usually strongly rooted. However, as we stated in chapter 3, both girls had average or above average marks from physics. Therefore, we can consider them students with typical non-Newtonian conceptions, who have at least average general learning abilities. As we will discuss later, their general strategy of solving the tasks is similar to the HPK students.

We can split solving of a multiple-choice task into three cognitive processes: general strategy of solving the task/the decision-making process, stem/task interpretation and options interpretation.

General strategy & stem-options transitions
As for general strategy, both girls first read the whole stem properly and then they visited options and returned to the stem. The strategy allowed us to separate the analysis into reading the stem and choosing the answer. This strategy shows organized approach to solving the task, which can indicate, to some extent, motivation to deal with the task and general knowledge about the multiple-choice task format. The incorrect answer seems to arise from

the lack of physics knowledge rather than from lack of general skills. Only six out of 35 students solved the two tasks correctly. All of them used the same general strategy as the incorrectly answering students in our study. Three of the correctly answering students (similarly to Jane) made a short look at one option before completing the stem reading.

All the correctly answering students provide about three transitions from the stem to the options (specifically: 4, 2, 3, 4, 2 and 3). Eve showed similar behavior. So did Jane, but only during solving the second task. In the first task, she made 16 transitions from the stem to the options after reading the whole stem and during her decision-making process. Jane's approach is in line with findings of Lindner et al. (2014) that show that LPK students demonstrated "overview behavior" more often and inspected and revisited the stem more frequently, while HPK students tended to arrive at their answer more quickly. In line with the gaze bias effect, they tend to show more "focused behavior" then LPK students.

When solving the first task, Jane distributed her attention among the options more equally and she paid the least amount of attention to the option that she finally chose as her answer. When solving the second task, her strategy showed a more focused behavior typical for HPK students. Firstly, she omitted two options at the beginning and, in the second half of her decision-making process, paid more attention to the option which she in the end selected as her answer. Eve showed a similar general strategy as HPK students when solving the first task: she made only several transitions between the stem and options and she paid the most attention to the option which she chose as her answer, albeit the incorrect one. In the second task, she was probably influenced by the first task and paid a lot of attention to the option which she previously chose as her answer. After rereading the stem and her correction of the task, she left the option completely and paid almost all attention to the option which she in the end chose as her answer. Typically, we observe different strategies according to the students' level of knowledge (HPK and LPK students). As we have shown above, students with typical non-Newtonian concepts solving the tasks incorrectly show strategies which are essentially similar to strategies of HPK students. This may indicate that typical non-Newtonian concepts are indeed alternative ideas rather than an absence of a concept relating to a given topic. Further research in education field could focus more on the problem of alternative concepts.

Five of six students, who solved the tasks correctly, arrived at their solution quickly. In three cases, they fixated the most on two options (including the correct one). In one case, a student paid much more fixations to the correct option, whilst another student did the contrary. One correctly answering student was not so confident and went through all the options with interest.

Though Eve solved both tasks incorrectly, her general approach is more similar to the approach of successful students than Jane's.

Stem/task interpretation

Jane tended to spend much more time on reading the stem in contrast to correctly answering students, while Eve did not. Further analysis of reading the stem consists mainly of identification of longer fixations and observation of regressive saccades. Average fixation duration was an appropriate threshold for identifying longer fixations. Both girls focused these fixations on the key terms of the stem. But Eve did not make a longer fixation on the word "constant" in the "rocket before" task. As for the regressive part of the saccade, we could easily identify in-word and between-words saccades. These are presented in the graphic visualisation of saccadic movements in Fig. 11. Moreover, we observed an unusual type of in-word saccade for Eve, when she sometimes did short vertical saccades in the middle of the key word. Probably, she worked with the word or the previous key terms in her mind. We did not identify many other returns during the reading. Those who solved tasks correctly show very similar way of reading the stem with emphasis (either by longer fixation duration or regressive saccades) on the key terms as well.

Options interpretation

Jane paid attention to the additional, decorative picture and in both tasks she showed error similar to "graph as a picture error" when she considered trajectory of the rocket (line) as a horizontal line or "constant" speed. She strongly shows a typical wrong idea ignoring Newton´s First Law, thinking that when the only acting force stops acting on the object, the object has to decrease its speed. Thanks to the eye-tracking data we could observe it during the first task as well, even though only the second task aimed at understanding of this law.

Though Eve was more interested in three of the options in the first task, her attention was focused on the word "constant" in all cases. If we relied solely on summarized AOI data, this finding would go undetected. Though she was interested in two options in the second task, she later realized what we the task was actually about and then her choice was quite confident and quick. When she started considering one option as her answer, it was evident by a much longer fixation.

Four of the six students who solved the "rocket after" task correctly were not as confident about their correct answer and paid very much attention to the option "decreasing" or "constant and decreasing", which indicates they also needed to think about the typical non-Newtonian conception. One student even followed the decreasing line of the graph in the option with his eyes, which he did it not do for the other graphs.

Limits of the study

Presented tasks are in line with the FCI test (Hestenes et al., 1992), which is considered a quick-answer test, i.e. students typically make their choice very quickly. The eye-tracking method allowed us to follow students' way of thinking to some extent. However, follow-up introspective verbal protocol when students watched and commented on the animated gaze plots did not provide us with more additional information as students simply did not know what they should say more about it. An in-depth interview about all the tasks focused on the participants' understanding of the two Newton's Laws would be beneficial.

We focused heavily on students' reading of the stems. We could observe whether they focused on relevant parts of the stem and did not miss important information needed for solving the task successfully. We can think of reading the stems as reading a specialized text. Including reading of a casual text or solving a task from a different field into the study design could allow us to interpret the stems reading in a more contextualized way.

7   Conclusion

We used the eye-tracking method to qualitatively analyse the eye movements of two low know-ledged (LPK) students who solved multiple-choice physics test. The test was focused on understanding of Newton's 1st and 2nd Law. Both

girls achieved the same test score (zero) and chose the same incorrect options. In the paper-pencil version of the test they would have received the same results. Previous studies (e.g. Tai et al., 2006) focused on qualitative analysis of scanpath differences between HPK and LPK students. Our qualitative analysis revealed differences in approach to solving the problems between LPK students as well. Moreover, we also showed that summarized AOI data sometimes do not provide enough information for proper interpretation if we create the AOIs too large (e. g. each option as a separate AOI). Finally, we presented several ways how to visualize eye-tracking data of a multiple-choice task solver including visualisation of reader's data.

Further research recommendations

All this information is available after approximately 10 minutes of interaction with a student. Solving of each task took approximately 1 minute. In school practice, such kind of data could help a teacher to start individual dialogue with each student. For example, Eve was very strongly convinced about her answer, so the teacher could start a dialogue about these incorrect ideas and conceptions. Jane struggled with understanding the task itself. She needed to read the stem repeatedly, she would probably need further explanation or rephrasing of some of the terms, and visual clues. It would also be helpful for Jane to talk about the "graph as a picture error".

However, eye-trackers being commonly available in schools is still a distant future. Nevertheless, we can conduct research, which can impact the school practice now or later. Firstly, as we showed, focus on incorrectly answering students and observing their decision-making process even when they solve such tasks is valuable for deeper understanding of different students' conceptual understanding. Secondly, as highlighted by Eve's rare vertical saccades present probably when she was thinking about important gained information, focus on students' eye movements when they are not extracting information but thinking during the text processing seems equally reasonable as a theme for further research, as does observing students while they read specialized text. Results of these studies can further serve as a background for teachers when they would like to work with eye-trackers directly in the school practicr or for creators of anticipating educational texts, such as in the HyperMind project (HyperMind project, 2018).

References

Bax, S. (2013). *Readers' cognitive processes during IELTS reading tests: Evidence from eye tracking.* ELT Research Papers 13-06. London: British Council. ISBN 978-0-86355-720-0

Beichner, R., J. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics*, *62*(8), 750–762.

Blascheck, T., Raschke, M., & Ertl, T. (2013). Circular heat map transition diagram. In *Proceedings of the 2013 Conference on Eye Tracking South Africa* (pp. 58–61). New York: ACM.

Blascheck, T., Kurzhals, K., Raschke, M., Burch, M., Weiskopf, D., & Ertl. T. (2014). State-of-the-art of visualization for eye tracking data. In R. Borgo, R. R. Maciejewski, & I. Viola (Eds.), *Proceedings of Eurographics Conference on Visualization (EuroVis).*

Bojko, A. (2013). *Eye tracking the user experience: A practical guide to research.* New York: Rosenfeld Media.

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A., & Kliegl, R. (2011). Coregistration of eye movements and eeg in natural reading: Analyses and review. *Journal of Experimental Psychology: General*, *140* (4), 552–572.

Dunn, R., Dunn, K., & Price, G. E. (1989). *Learning style inventory (LSI): Price Systems*, Incorporated (PO Box 1818, Lawrence 66044).

Gegenfurtner, A., Lehtinen, E., & Saljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, *23*(4), 523–552.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher*, *30*(3), 141–158.

HyperMind project. Websites. Available from https://uedu.uni-kl.de/arbeitsfelder/arbeitsfeld-01/hypermind/english

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van Dee Weijer, J. (2011). *Eye tracking.* Oxford: Oxford university press.

Kekule, M., & Viiri, J. (2018). Students' approaches to solving R-FCI tasks observed by eye-tracking method. *Scientia in educatione*, *9*(2), 117–130.

Kekule, M., & Krejčí, A. (2019, in print). Žákovské čtení textu z učebnice sledované oční kamerou a role otázek při porozumění tomuto textu. In *Moderní trendy v přípravě učitelů fyziky 2019*.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General*, *35*(1), 12.

Kozhevnikov, M., Motes, M., & Hegarthy, M. (2007). Spatial visualization in physics problem solving. *Cognitive Science*, *31*(4), 549–579.

Kübler, T., Sippel, K., Fuhl, W., Schievelbein, G., Aufreiter, J., Rosenberg, R., Rosenstiel, W., & Kasneci, E. (2015). Analysis of eye movements with Eyetrace, In *International Joint Conference on Biomedical Engineering Systems and Technologies* (pp. 458–471). Cham: Springer.

Kunze, K., Kawaichi, H., Yoshimura, K., & Kise, K. (2013a). Towards inferring language expertise using eye tracking. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 217–222).

Kunze, K., Kawaichi, H., Yoshimura, K., & Kise, K. (2013b). The Wordometer -- estimating the number of words read using document image retrieval and mobile eye tracking. In *12th International Conference on Document Analysis and Recognition*, IEEE, Washington, DC, USA.

Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., …. Tsai, C. C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, *10*, 90–115.

Lindner, M. A., Eitel, A., Thoma, G.-B., Dalehefte, I. M., Ihme, J. M., & Köller, O. (2014). Tracking the decision-making process in multiple-choice assessment: Evidence from eye movements. *Applied Cognitive Psychology*, *28*(5), 738–752.

Madsen, A. M., Larson, A. M., Loschky, L. C., & Sanjay Rebello, N. (2012). Difference in visual attention between those who correctly and incorrectly answer physics problems. *Physical Review Special Topics-Physics Education Research*, *8*(1), 010122.

Mareš, J., & Slavík, V. (1990). *Dotazník stylu učení.* Praha: Institut pedagogicko-psychologického poradenství České republiky.

Nieminen, P., Savinainen, A., & Viiri, J. (2010). Force Concept Inventory-based multiple-choice test for investigating students' representational consistency. *Physical Review Special Topics – Physics Education Research*, *6*, 020109.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422.

Rayner, K. & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs: Prentice Hall.

Reichle, E., Warren, T., & McConnell, K. (2009). Using E-Z reader to model the effects of higher-level language processing on eye movements during reading. *Psychonomic Bulletin & Review, 16*(1), 1–20.

Šmideková, Z. (2018). Eye-tracking v prostredí školskej triedy, *Pedagogika*, *68*(1), 25–50.

Tsai, M.-J., Hou, H.-T., Lai, M.-L., Liu, W.-Y., & Yang, F.-Y. (2012). Visual attention for solving multiple-choice science problem: An eye-tracking analysis. *Computers & Education*, *58*, 375–385.

Tai, R. H., Loehr, J. F., & Brigham, F. J. (2006). An exploration of the use of eye-gaze tracking to study problem-solving on standardized science assessments. *International Journal of Research & Method in Education*, *29*(2), 185–208

Viiri, J., Kekule, M., Isoniemi J., & Hautala, J. (2017). Eye-tracking the effect of representation on students' problem solving approaches. In *Proceedings of the annual symposium of the FMSERA 2016* (88–98). The Finnish Mathematics and Science Education Research Association.

Voßkühler, A., Nordmeier, V., Kuchinke, L., & Jacobs, A.M. (2008). OGAMA – OpenGazeAndMouseAnalyzer: Open source software designed to analyze eye and mouse movements in slideshow study designs. *Behavior Research Methods*, *40*(4), 1150–1162.

Authors

Martina Kekule, Ph. D., Charles University, Faculty of Mathematics and Physics, Department of Physics Education, V Holešovičkách 2, Prague 180 00, Czech Republic
email: martina.kekule@seznam.cz

Iva Špaňová, M. Sc., Základní škola, Dobruška, Pulická 378, okres Rychnov nad Kněžnou
Pulická 378, 51801 Dobruška, Czech Republic
email: iva.spanova@email.cz

Jouni Viiri, Ph. D. University of Jyväskylä, Department of Teacher Education,
Alvar Aallon katu 9, 40140 Jyväskylä, Finland
email: jouni.p.t.viiri@jyu.fi

Výhody využití metody oční kamery při sledování žákovských kognitivních procesů při řešení testů s výběrem odpovědi

Abstrakt: Metoda oční kamery se v rámci různých studií ukázala jako vhodnou pro studování kognitivních procesů žáků. V článku se zaměříme na výhody použití metody sledováním žáků středních škol při řešení testu s otázkami s výběrem odpovědi. Konkrétně se zaměříme na kvalitativní analýzu dat týkajících se pohybů očí dvou žákyň řešících úlohy z fyziky. Obě žákyně řešily všechny testové úlohy nesprávně, obě navíc vybraly jako odpověď stejné alternativy. V případě klasického zadání testu (vyplňování papírových dotazníků), bychom dostali od obou žákyň naprosto stejný výstup. Data z oční kamery nám však ukázala jak podobné přístupy řešení, tak i důležité odlišnosti, na základě kterých by pak vyučující pracovali s každou dívkou odlišně. Obě dívky četly zadání pozorně a věnovaly důraz klíčovým pojmům. Je tedy zřejmé, že špatná volba odpovědi nejde na vrub nedostatečného čtení zadání. Obě dívky zřejmě mají typickou miskoncepci týkající se 1. Newtonova zákona. Liší se zejména v jejich jistotě výběru odpovědi. Eva byla v odpovídání rychlejší a zabývala se zejména možností, kterou nakonec vybrala jako odpověď. Naproti tomu Jana se velmi často vracela k zadání, často nejprve přemýšlela o odpovědi, kterou nakonec nevybrala jako správnou. U Evy jsme pozorovali netypické vertikální sakády nad vybranými klíčovými pojmy. Jana při pohledu na spíše ilustrativní obrázek projevila typickou miskoncepci "graf jako obrázek". Detailní analýzou jsme ukázali, že souhrnná data na tzv. oblast zájmu mohou být zavádějící, pokud je tato oblast příliš velká (např. jedna alternativa v testové otázce). Dále jsme představili možné indikátory pro kvalitativní analýzu a možnosti vizualizace zpětných sakád.

Klíčová slova: oční kamera, fyzikální vzdělávání, zpětné sakády, test s výběrem odpovědi, řešení problémů

Appendix I

Task Rocket before.

Test version in the Czech language.

Raketa letí ve vesmíru, kde na ni nepůsobí žádné vnější síly. V čase t_0 je zapnut raketový motor, který působí na raketu konstantní silou ve směru pohybu. V čase t_8 je motor vypnut.

Během časového intervalu $t_0 - t_8$ je velikost rychlosti rakety:

a) zpočátku konstantní a poté se snižuje.

b) zpočátku se zvyšuje a poté je konstantní.

c) neustále se snižuje.

d) neustále se zvyšuje.

e) konstantní.

Translation into English:

A rocket drifts sideways in outer space. The rocket is subject to no outside forces. At the instant of time t0 the rocket's engine is turned on and produces (end of the 1st row in the Czech)

a constant force on the rocket. The force is in the direction of the motion. At the instant time t8 the rocket's engine is turned off. (end of the 2nd row in the Czech)

During the time interval t_0-t_8 the speed of the rocket is:

a) constant for a while and decreasing thereafter.

b) increasing for a while and constant thereafter.

c) continuously decreasing.

d) continuously increasing.

e) constant.

Task Rocket after.

Test version in the Czech language.

Raketa letí ve vesmíru, kde na ni nepůsobí žádné vnější síly. V čase t_0 je zapnut raketový motor, který působí na raketu

konstantní silou ve směru pohybu. V čase t_8 je motor vypnut.

Který z následujících grafů nejlépe popisuje závislost velikosti rychlosti rakety na čase po okamžiku t_8?

a) rychlost — čas — t_8

b) rychlost — čas — t_8

c) rychlost — čas — t_8

d) rychlost — čas — t_8

e) rychlost — čas — t_8
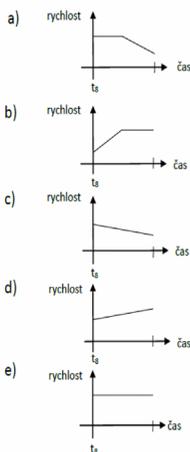
Translation into English:

A rocket drifts sideways in outer space. The rocket is subject to no outside forces. At the instant of time t_0 the rocket's engine is turned on and produces (end of the 1st row in the Czech)

a constant force on the rocket. The force is in the direction of the motion. At the instant time t8 the rocket's engine is turned off. (end of the 2nd row in the Czech)

Which of the following alternatives best describes the speed of the rocket after the time t_8?

For options a) – e) see graphs (speed (time)) above.

Appendix II

"Rocket before" task. Single fixation duration (vertical axis in ms) during the first reading of the stem. Values less than 60 ms were omitted. The vertical line shows the average fixations duration, which can serve as a reasonable threshold.

Jane
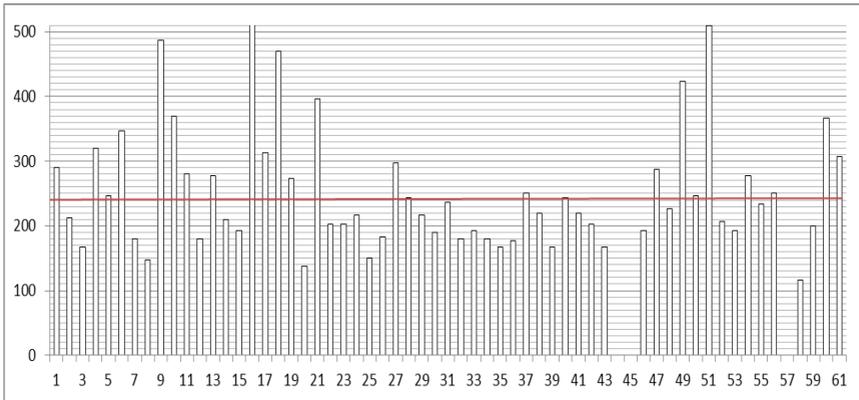
Eve
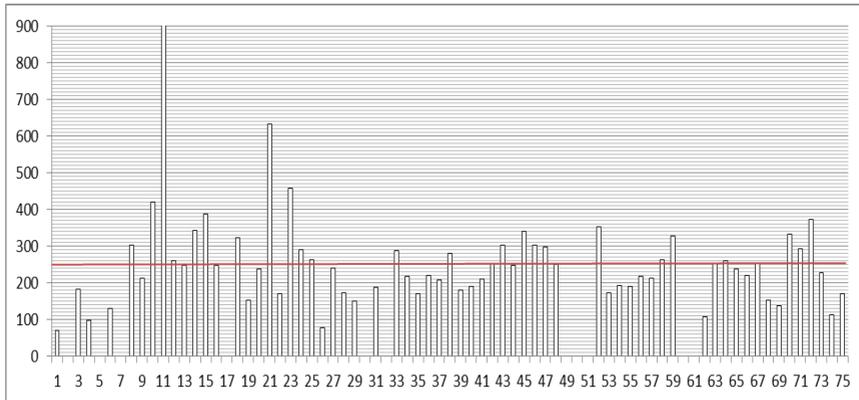
Appendix III

Gaze plots for two participants Jane and Eve and two tasks. The longer fixation duration, the larger the circle.

"Rocket before" task

"Rocket after" task