The Effect of Subject Familiarity on Comprehension and Eye Movements during Reading

Leana Copeland
Research School of Computer Science
Australian National University
leana.copeland@anu.edu.au

Tom Gedeon
Research School of Computer Science
Australian National University
tom.gedeon@anu.edu.au

ABSTRACT
We investigate factors affecting reading and overall comprehension of the underlying meaning and concepts within a piece of text using eye movements. Our objective is to identify eye movement measures that will predict reading comprehension, and intend to apply them in eLearning to create dynamic learning environments that can use eye movement to detect reader comprehension. We found that the self-reported familiarity of readers with the subject of documents affects their reading behaviour but not their total comprehension score, and found that we could identify answer-seeking behaviour and a measure of their actual familiarity with the text content using eye gaze.

Author Keywords
Eye Movement measures; eye gaze; assessing reading comprehension; reading analysis

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

We investigate factors affecting reading and overall comprehension of the underlying meaning and concepts within a piece of text using eye movements. We seek to find eye movement measures that can be used to quantify reading comprehension. The eventual use of such measures is to create dynamic learning environments that provide feedback and dynamic material based on personal comprehension levels. We begin by looking at the situation where participants are required to read tutorial content and answer comprehension questions. We investigate if there are eye movement measures that can be used to identify a distinctive reading behaviour that we have termed answer-seeking behaviour. This is part of on-going research broadly aimed at detecting when a student is finding it difficult to understand material and to help mediate efficient learning.

EYE MOVEMENTS DURING READING

Eye movements can be broadly characterised as fixations and saccades. A fixation is where the eye remains relatively still to take in visual information. A saccade is a rapid movement that transports the eye to another fixation. The reason for this behaviour is due to the anatomy of the eye. At the centre of the retina is a special part of the eye that sees in fine detail called the fovea. The foveal region of the eye is very small, being only about 0.2mm in diameter. Around the point of fixation visual acuity extends about 2° (Rayner, 1998). This means that humans see very little in detail at any fixation and is why the phenomena of saccades are observed. The eye must move around rapidly so that it can compose a more detailed view of the environment.

Generally when reading English, fixation duration is around 200-300 milliseconds, with a range of 100-500 milliseconds and saccadic movement is between 1 and 15 characters with an average of 7-9 characters (Rayner, 1998). The majority of saccades are to transport the eye forward in the text when reading English, however, a proficient reader exhibits backward saccades to previously read words or lines about 10-15% of the time (Rayner, 1998). These backward saccades are termed regressions. Short regressions can occur within words or a few words back and may be due to problems in processing the currently fixated word, overshoots in saccades, or oculomotor errors (Rayner, 1998). Longer regressions occur because of comprehension difficulties, so the reader tends to send their eyes back to the part of the text that caused the difficulty (Rayner, 1998).

Eye Movements and Reading Comprehension

Eye movements can be used to understand the ongoing cognitive processes that occur during reading (Rayner, 1998). Comprehension of text can have significant effects on the eye movements observed (Rayner et al, 2006). A number of studies have shown there are numerous variables based around comprehension functions that can have influence on eye movements during reading. The variables include: semantic relationships between words, anaphora and co-reference, lexical ambiguity, phonological ambiguity, discourse factors, stylistic conventions, and syntactic disambiguation (Rayner, 1998). These variables have different effects on eye movement, causing them to deviate from the default reading process. For example, "garden-path" sentences are syntactically ambiguous and induce regressions to resolve the comprehension problems (Frazier and Rayner, 1982). Eye movements have been shown to reflect text difficulty (Rayner et al., 2006).

Related Work

Eye gaze patterns can be used to detect what kind of task the participant is performing (Iqbal and Bailey, 2004) or

OzCHI'13, November 25 - 29 2013, Adelaide, Australia

whether a person is reading or not (Cambpell and Maglio, 2001) as well as if they are reading or skimming (Buscher et al., 2008). Furthermore, there are several applications that use eye gaze to provide reading assistance. These include iDict, a reading aid designed to help readers of a foreign language (Hyrskykari et al., 2000). iDict uses eye gaze to predict when a reader is having comprehension difficulties. If the user hesitates whilst reading a word then a translation of the word is provided along with a dictionary meaning. Another is The Reading Assistant (Sibert et al., 2000) that uses eye gaze to predict failure to recognise a word. The Reading Assistant then provides auditory pronunciation of the word to aid in reading. These applications work on the assumption that the user pauses on a problematic word, and then the systems provides feedback about that word. They do not look at overall text comprehension or provide feedback about the overall comprehension of that text.

METHOD

A user study was conducted to collect participants' eye gaze as they read a tutorial and completed a quiz based on the tutorial's content. The tutorial and quiz were coursework from a first year Computer Science course taken at the Australian National University. There were 15 (6 female, 9 male) participants aged between 17 and 31 who took part in the study. Of these 4 of the participants stated that English was not their first language. The choice of participants is based on the target user group of the eventual online learning environment, which is university students.

Participants read 9 screens of textual content each covering a specific area about the main topic of the tutorial (Web Search). Each screen was 400 words long. The tutorial content was accessible via the online learning environment used at ANU, called Wattle (a Moodle variant). After each screen, participants were required to answer two questions to measure their comprehension (18 questions in total); one of the questions was multiple-choice and the other was cloze (fill-in-the-blanks). These two types of questions were used because they can assess different forms of comprehension (Fletcher, 2006). When presented with the questions, participants were also given the opportunity for a second read-through of the content to aid in answering questions.

The study was displayed on a 1280x1024 pixel Dell monitor. Eye gaze data was recorded at 60Hz using Seeing Machines FaceLAB 5 infrared cameras mounted at the base of the monitor. The study involved a 9-point calibration prior to data collection for each participant. As the data recorded is a series of gaze points, EyeWorks Analyze was used to pre-process the data to give fixation points. The parameters used for this were a minimum duration of 75 milliseconds and a threshold of 5 pixels.

RESULTS AND DISCUSSION

Effect of Time on Score

On average it took participants 34 minutes to complete the tutorial and quiz with a range of 18 to 50 minutes. We found a moderate negative correlation (r=-0.4) between time taken and final grade. As time taken to complete the quiz increased the comprehension score is seen to decrease. Familiarity, or prior knowledge, of the subject matter could account for this relationship.

Effect of Familiarity

Participants were asked to subjectively rate their familiarity with the topic of the content. The participants were grouped into "familiar" (4 participants), "somewhat familiar" (7 participants), and "not familiar" (4 participants). On average the "familiar" group received a total mark of 15.5, the "somewhat familiar" group achieved an average score of 17, and the "not familiar" group received 15.1. The differences in average total comprehension mark between the three groups were not statistically significant (using three two-sided unpaired Student's t-test with significance level of $p<0.05$).

On average the participants in the "familiar" group completed the tutorial in 30 minutes, the participants in the "somewhat familiar" group completed it in 35 minutes on average, and those in the "not familiar" group completed it in 38 minutes on average. These differences were found to be statistically insignificant (using three unpaired two-sided Student's t-tests with significance level of $p<0.05$). This indicates that the relationship between time taken to complete the quiz and total score obtained is not significantly due to subjective familiarity with the content. Instead, we observed that increased time taken to complete the quiz and a decrease in total mark is more likely indicative of (1) how difficult the participants found the material and questions and (2) how confident the participants were answering the questions.

The eye movement measures calculated for the participants within the familiarity groups are shown to be different. The same text is shown to the participants twice, first to read through (measures detailed in Table 1) and a second time to use in answering each of the questions (measures detailed in Table 2). Considering the eye movements for these two presentations of text read can reveal the effect that perceived familiarity has on how the first presentation of the content is read, as well as on how the text is subsequently used to answer the questions. When reading the text the first time, the three groups can be differentiated statistically based on mean fixation duration as well as total fixation duration (using unpaired two sided Student's t-test with significance level of $p<0.05$). The participants in the "familiar" group had the lowest average number of fixations and the longest total fixation time, followed by the "somewhat familiar" group. The participants who were least familiar displayed the most fixations of the shortest duration. The differences in the average numbers of fixations recorded for each of the groups is not an indication that one of the groups is reading the text more thoroughly compared to the other. Rather, that the participants within each of the groups appear to read the text differently.

When considering the differences of the categories of text, the greatest difference in reading behaviours can be seen when comparing the eye movements observed for the second read-through. On average the participants in

the "familiar" group displayed 77 fixations for on average 0.19 seconds and 52% of the fixation transitions were regressions. The participants in the "not familiar" group exhibited 199 fixations for an average of 0.15 seconds and 42% of the fixation transitions were regressions. These differences were found to be statistically quite different (using unpaired two sided Student's t-test with significance level $p<0.01$). The participants in the "somewhat familiar" group also had fewer fixations, for longer mean duration and higher regression ratio compared to that observed from the participants in the "not familiar" group. These differences were found to be statistically different (using unpaired two sided Student's t-test with significance level $p<0.01$).

	Average for Group		
Eye Movement Measure	**Familiar**	**Somewhat Familiar**	**Not Familiar**
Number of Fixation	242.1	271.3	255.9
Mean Fixation Duration	0.23	0.18	0.17
Total Fixation Time	60.6	51.3	43.2
Number of Regressions	73.3	85.3	66.7
Regression Ratio	0.32	0.33	0.30
Ave. Regression Length	283.9	267.0	284.2
Ave. Forward Saccade Length	99.4	102.2	91.4

Table 1. Average eye movement measures by familiarity group for eye movements recorded for the first read through of the content.

	Average for Group		
Eye Movement Measure	**Familiar**	**Somewhat Familiar**	**Not Familiar**
Number of Fixation	77.1	104.8	198.6
Mean Fixation Duration	0.19	0.16	0.15
Total Fixation Time	15.3	18.8	32.8
Number of Regressions	34.8	44.2	76.3
Regression Ratio	0.52	0.47	0.42
Ave. Regression Length	244.2	234.0	260.7
Ave. Forward Saccade Length	132.1	141.7	127.6

Table 2. Average eye movement measures by familiarity group for eye movements recorded for the second read through of the content.

It has been shown that when text is difficult for the reader to understand there is an increased processing time, and regressive eye movements are increased (Rayner et al, 2006). There are clear differences between the groups in the reading time and regressive behaviour seen for the second read through. Given that all three groups have fairly consistent reading time and regressive eye movements for the first read through it is reasonable to conclude that the text itself was not hard to read. Instead the time taken to read the text on the second reading was due to difficulties answering the questions based on the text.

Additionally, there is no correlation between the number of fixations seen for the first read-through of the text to the second read-through of the text. Having only a few fixations (potentially skimming) the content the first time did not correlate to more fixations (deep reading) observed for reading the content the second time, and comprehensively reading the content the first time did not mean spending less time reading the content the second time.

Defining Answer-Seeking Behaviour

There is little correlation ($r=-0.2$) between the participants' average number of fixations recorded for the reading the content the second time and the total score that participant received. What this would suggest is that there are other factors that are associated with answering the questions correctly. To measure this we will consider the reading behaviour during the first and second read through of the content. In the first read through participants are simply required to read the content with no objective to accomplish other than some unknown questions to be asked later. When presented with the content for the second time the objective is clear and the reading behaviour observed is what we have termed answer-seeking behaviour. This is a new measure of reading comprehension for the purpose at determining overall text comprehension. The difference seen in this behaviour can be used to make inferences about the participant's confidence in answering the questions relating to the content. Here confidence is defined as the participant's prior and current knowledge as well as the participant's certainty in their answers and their ability to answer the question correctly. There will be participants who do not know the answers and need to find them, there will be others who think they know the answer and want to double check that it is right, and there will be those that think they have the right answer but do not double check. In any case there is a range of confidence described in answering of the questions, which can be associated with the understanding of the material, the subjective familiarity with the subject and self-assurance in having the right answer.

We have shown that the number of fixations and total fixation time recorded for the reading of the second display of the content reflect a measure of answer-seeking behaviour. In this case the participants in the "not familiar" group displayed the most answer-seeking behaviour followed by the participants in the "somewhat familiar" group and finally those in the "familiar" group showed the lowest amount of answer-seeking behaviour (see Table 2). Using the same eye movement measures all three groups read the content for the first time with the same intensity (see Table 1) so this difference is not a

consequence of lack of reading attention to the content the first time.

Answer-seeking behaviour does not guarantee that the correct answer is given and neither does the lack of answer-seeking behaviour. Within the group of participants that were "not familiar" with the content, there is a strong positive correlation (r=0.9) between the average number of fixations observed when reading the second display of the content and the score they received for the quiz. For the participants in the "somewhat familiar" and "familiar" groups the opposite was observed, there is a moderate negative correlation (r=-0.7 and r=-0.5, respectively) between the average number of fixations for the second display of the content to the total score. Therefore, more fixations here would be associated with lower certainty in answering and finding the answer to the questions.

While there is no strong correlation between fixations and perceived familiarity across all three groups, we were able to observe distinctively different reading behaviours in participants with little subject familiarity and those with some or a high degree of familiarity. The participants who had no familiarity with the subject did better when they spent longer re-reading the second display of the text, taking a thorough approach and confirming that they had the correct answer. On the other hand the participants with some familiarity with the subject appeared to show the opposite effect, whereby a longer and more considered reading approach was more indicative of lower confidence in answering the question.

CONCLUSIONS AND FURTHER WORK

In this study we considered the effect of subjective familiarity with a topic on participants' performance on a tutorial and quiz and their eye movements. Whilst on average these groups performed the same on the quiz, there are differences in the eye movements observed for these groups of participants. Most importantly these differences are seen in the reading behaviour recorded for the second display of the content, which we have termed answer-seeking behaviour. This is a new measure that we propose to use to evaluate overall text comprehension. Overall text comprehension has not been assessed in this way to date for the purpose of developing a HCI tool. We have shown that the number of fixations and total fixation duration can be used to determine the relative familiarity as well as the extent to which answer-seeking is being performed.

The outcomes of this current study are aimed at being able to detect a student's reading behaviour and understanding of text. This data will be analysed further to assess how the eye movement measures can be related back to the participants' understanding of the content and the questions. We will investigate further the use of answer-seeking behaviour as a measure of overall reading comprehension as well as reliable and effective ways of measuring it. This is the first step in identifying differences in comprehension processes and formulating ways to evaluate them using eye movements as an implicit comprehension measure.

The purpose of this work is to ultimately implement the use of reading comprehension measures in an adaptive eLearning environment. In such a system, students can be presented with more challenging content if they are not being challenged by the normal level of content due to prior knowledge or ease in understanding the concepts. On the other hand reduced technical information can be presented to students who are struggling with the normal level content. Furthermore, information about the nature of how students are reading content can be more useful than their ability to answer the questions on the content. Instructors could use the eye movement behaviour to identify parts of subject content/teaching material that may be too technical, ambiguous, or not well structured. Using this information they can then restructure the content to optimise the learning experience for their students. Implementation of such a system would require eye-tracking capabilities with an interface that is designed to calculate reading comprehension measures in real time and provide feedback based on the values of the measures calculated. Development of such a system is part of ongoing research.

REFERENCES

Buscher, G., Dengel, A., and Van Elst, L. Eye movements as implicit relevance feedback, CHI '08 Extended Abstracts on Human Factors in Computing Systems (2008), 2991-2996.

Campbell, C.S. and Maglio, P.P A robust algorithm for reading detection. Proceedings of the 2001 workshop on Perceptive user interfaces (2001) ACM.

Fletcher, J.M. Measuring Reading Comprehension. Scientific Studies of Reading, 10 (2006), 323-330.

Hyrskykari, A., Majaranta, P., Aaltonen, A. & Räihä, K.-J. Design issues of iDICT: a gaze-assisted translation aid. Proceedings of the 2000 symposium on Eye tracking research & applications (2000), 9-14.

Iqbal, S.T. and Bailey, B.P. "Using Eye Gaze Patterns to Identify User Tasks. Proceedings of The Grace Hopper Celebration of Women in Computing, (2004).

Rayner, K. Eye Movements in Reading and Information Processing: 20 Years of Research. Psychological Bulletin, 124, 3 (1998), 372-422.

Rayner, K., Chace, K.H., Slattery, T.J., and Ashby, J. Eye Movements as Reflections of Comprehension Processes in Reading. Scientific Studies of Reading 10, 3 (2006), 241-255.

Sibert, J. L., Gokturk, M. & Lavine, R. A. The Reading Assistant: Eye Gaze Triggered Auditory Prompting for Reading Remediation. Proceedings of the 13th annual ACM symposium on User interface software and technology (2000), 101-107.