# Measuring Reading Comprehension using Eye Movements

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

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

***Abstract*—We investigate eye movement measures and methods for predicting reading comprehension. This builds on previous work on factors affecting reading comprehension, namely perceived familiarity with the text content. We further investigate answer-seeking behavior and present a method for measuring and comparing this behavior. The number of fixations, number of regressions, and total fixation time are an indicator of reading intensity and the intensity of reading is related to comprehension. We show that a feed-forward backpropagation neural network can be used to predict subjective comprehension scores as well as quiz scores. We propose using the degree of answer-seeking behavior to measure how question difficulty and as an implicit measure of how difficult a participant finds a tutorial and quiz. Such information is beneficial to apply in eLearning to create dynamic learning environments that use eye movement to predict implicit question difficulty as well as individual participant difficulty.**

**Keywords—Reading Analysis; Reading Comprehension; Eye gaze; adaptive eLearning environments; Reading Behavior**

## I. INTRODUCTION

We investigate measures for analyzing reading behavior 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. The intention is to apply them in eLearning to create dynamic learning environments that use eye movement to detect reader comprehension. We begin by looking at the situation where students read tutorial content and answer comprehension questions. We investigate methods for identifying and comparing answer-seeking behavior. Further, we explore how inferences about reading behavior can be used in an online learning environment to optimize learning. This is part of on-going research that will be more broadly aimed at detecting when a student is having difficultly understanding material to ameliorate effective learning. We aim to increase learning, a fundamental human cognitive process, through interactive communication with the learning environment. We therefore contribute to CogInfoCom research by utilizing knowledge about human cognition to develop information transfer techniques between humans and infocommunications systems that are primarily based on Internet technology and for the purpose of education.

We first review the literature on reading analysis using eye movements and how previous work can be applied to inferring comprehension. Then the results of a user study will be presented and discussed. Finally, we will conclude and indicate how these results will be used in the future work of developing an adaptive eLearning environment.

### A. Eye Movements During Reading

Eye movements can be broadly characterized 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. 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° [1].

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 [1]. 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 [1]. 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. However, longer regressions occur due to comprehension difficulties, as the reader tends to send their eyes back to the part of the text that caused the difficulty [1].

### B. Reading Comprehension and Eye Movements

Reading comprehension is a skill that must be taught and requires constant education. When understanding language, we integrate ideas in the text and form a mental model that is an abstraction of the conglomeration of ideas [2,3]. The actual text read is not remembered verbatim, it is the ideas and constructed representation that are [2,3]. Comprehension of the text can have significant effects on the eye movements observed [1,4]. Studies have shown there are numerous variables that influence eye movements during reading: semantic relationships between words, anaphora and co-reference, lexical ambiguity, phonological ambiguity, discourse factors and stylistic conventions, and syntactic disambiguation. For example, garden-path sentences are syntactically ambiguous and induce regressions to resolve the

978-1-4799-1546-0/13/$31.00 ©2013 IEEE

comprehension problems [5]. Eye movements have also been shown to reflect text difficulty [4].

### C. Related Work

Eye gaze patterns can be used to detect what kind of task the participant is performing [6,7] or whether a person is reading or not [8,9] as well as if they are reading or skimming [10]. Previous studies have shown that even within the activity of reading, eye gaze patterns can be used to differentiate when individuals are reading different types of content [11].

One method of using eye gaze in human-computer interaction is to integrate eye gaze implicitly into the use of an interface. In such a system the user may not even be aware that their eye gaze is being used to control or alter the use of the system. An example of such a system is eye gaze based rendering. This is where the high-resolution image is displayed at the point of the users fixation [12]. This type of display exploits the fact that fine detail vision occurs only in the fovea and so only high resolution of image display is necessary at the point of fixation. Another example is the use of eye gaze to provide feedback about behavior such as in [13] where eye gaze is recorded to give implicit perceived relevance of pieces of text in a document. There are several applications that are used in reading assistance. iDict is a reading aid designed to help readers of a foreign language [14]. 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. This is somewhat similar to The Reading Assistant [15], which uses eye gaze to predict failure to recognize a word. The Reading Assistant then provides auditory pronunciation of the word to aid in reading.

## II. Our Prior Work

In previous work we analyzed the effect of perceived familiarity with the topic of the tutorial [16]. We found a relationship between perceived familiarity and answering behavior. In the first instance that participants are presented with the text there is no clear objective to the reading of the text other than knowing two questions will be asked about the content. In the second instance, the objective is clear as participants know the questions and can use the text as a reference to find the answers to the questions. This behavior of referencing the text to find the answer is what we have termed answer-seeking behavior. We used the number of fixations and total fixation time recorded for the reading of the second display of the content as indicators of answer-seeking behavior. We found that participants with less perceived familiarity displayed much greater answer-seeking behavior. This was also true for participants that stated they were familiar with the topic but did not answer questions correctly.

This current work is an extension of this research that endeavors to clearly define answer-seeking behavior as a metric that can be used to predict reading comprehension, give information about question and content difficulty, as well as provide a metric for implicit participant performance.

## III. 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 that took part in the study.

Participants read 9 slides of content covering a topic of the tutorial (Web Search). Each slide was 400 words long with an average Flesch Kincaid Grade Level of 12. All participants were university students and therefore had at least high school level education indicating that the readability of the slides should not be above their reading abilities. The tutorial content was accessible via the Wattle online learning environment at ANU (a variant of Moodle). After each slide participants answered two questions to measure their comprehension (18 questions in total); one question is multiple-choice and the other is cloze (fill-in-the-blanks). The two types of questions are to assess different forms of comprehension [17]. When presented with the questions, participants were also given the opportunity to re-read the content to aid in answering the questions. Once the participant finished the quiz and before shown their result, participants were asked to subjectively rate their overall comprehension on a scale of 1 to 10 with 10 being complete understanding.

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. EyeWorks was the software used to collect the data. The study involved 9-point calibration sequence.

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 0.075 seconds and a threshold of 5 pixels.

## IV. Analysis of Eye movement Measures for Comprehension

### A. Effect of First Read Through to Later Read Through

During the first presentation the participants read the content with no specific purpose. In the second presentation, participants know the questions and could use the text as a reference for answering the questions. The eye movements differed for the two readings. We hypothesized that there is a correlation between the two reading behaviors. Explicitly, the way a participant read the first display has some effect on how they read it the second time. We expected to see that if the participant read the content more thoroughly in the first presentation then he or she would spend less time reading the second presentation of content.

There is a medium negative correlation between both the number of fixations and the number of regressions ($r=-0.4$ for both measures), and the total fixation time ($r=-0.6$) observed in the first read through to the second read through. As more fixations and more regressions were observed in the first reading of the content, fewer fixations and regressions were

observed in the second read through. Furthermore, longer total fixation time observed for the first read through was also observed with a decrease in total fixation time observed for the second read through. This indicates that less reading of the second presentation is required to find the answers when more reading was done on the first presentation of the text. More reading is defined as higher numbers of fixations and regressions as well as a longer total fixation time. Whilst, more reading of the text does not guarantee understanding it can be noted that those participants who read the text more thoroughly the first time did not read the second presentation of the text as much.

The number of fixations, number of regressions, and total fixation time are an indicator of reading intensity. Reading intensity can be affected by many different factors such as how interested the reader is in the content, how hard the reader finds the content to understand, as well as external factors such as motivation for reading the content and answering the questions correctly. For these reasons reading intensity is not a direct indicator of reading comprehension. Instead it can be used as a supporting metric for measuring comprehension as we have shown that reading intensity of the first presentation of text is correlated to reading intensity of the second presentation of text. Reading of the second presentation of text is a measure of how confident a participant is in answering a question. More reading of the second presentation of text implies that the participant is not confident answering the question.

### *B. Further Exploration of Answer-Seeking Behaviour*

The areas of text each participant read can be distinguished into four categories; the first presentation of the slide content before questions are known, the multiple choice question, the cloze question, and the second read presentation of the slide content provided for assistance in answering the two questions. In this section we investigate eye movement measures for reading the four sections of text and their relationship to the outcome of answering the questions. When considering the eye movements observed for reading the multiple-choice questions there is a negative correlation between the number of fixations (r=-0.8), the total fixation time (r=-0.8), and the number of regressions (r=-0.7) to the score obtained for the multiple-choice question being read. Similarly for the cloze questions, the outcome of the question is correlated to the eye movements observed. There is a negative correlation between the number of fixations (r=-0.8), the maximum fixation duration (r=-0.8), the total fixation time (r=-0.8), the number of regressions (r=-0.8), and the regression ratio (r=-0.8) to the score obtained for the cloze question being read. Furthermore, there are also moderate negative correlations between eye movement measures observed for the reading of the second presentation of the content to the total marks received for that topic. The number of fixations (r=-0.5), the maximum fixation duration (r=-0.5), the total fixation time (r=-0.5), and the number of regressions (r=-0.5) are correlated to the score obtained for the both questions the section relates to.

From these correlations we observed that participants who read both the questions and the second presentation of the text more tended to do worse on the quiz. This observation is parallel to the previous finding that when questions are answered incorrectly more fixations for longer duration are observed. The definition of more reading is high numbers of fixations and regressions as well as a longer total fixation time. We can therefore deduce that more reading is indicative of the participant's lack of understanding of either the questions or the content. Time spent reading questions and referencing text for the questions is related to the participant's understanding whereby longer time spent answering the questions indicates less understanding.

Additionally, there are correlations between the eye movement measures observed when reading each type of question and the reading behavior seen when reading the second display of the content. We hypothesis that participants who re-read the question more are having difficulty answering the question and would therefore exhibit similar behavior when reading the second display of content. Indeed, we found positive correlations between the number of fixations observed for reading the multiple-choice question (r=0.7) and the cloze question (r=0.6) to the number of fixations observed for the second display of reading the content. Similarly, a positive correlation was found for total fixation time (r=0.7 and r=0.6, respectively) and number of regressions (r=0.8 and r=0.7, respectively) observed when comparing the eye movement recorded for the multiple-choice and cloze question to the second display of the content.

The participants who do not understand the question or the content well enough to answer the question seek to find the answer by re-reading both the question and the content. We term this answer-seeking behavior. Answer-seeking behavior is indicative of the participant's confidence in answering the questions. The participant's confidence is related to his or her actual understanding of the content, their perceived familiarity with the subject matter, as well as their confidence in his or her abilities to answer the questions correctly.

### *C. Defining Answer-Seeking Behaviour for Feedback*

We have established that answer-seeking behavior is an indicator of a participant's confidence in answering a question. Confidence is a product of many factors such as familiarity with the topic and is related to whether participants answer the questions correctly or incorrectly. More answer-seeking behavior indicates the participant has less confidence and is less likely to know the answer. It is beneficial to measure such behavior so that feedback can then be given based on the existence and extent of the answer seeking observed.

We propose measuring answer-seeking behavior by recording the large jumps between question and content regions and vice versa. Additionally, reading behavior within each region will be detected and recorded. The reading behavior is detected and recorded using the reading detection algorithm as defined by [10]. Both reading and skimming behavior is recorded as reading behavior in this analysis.

There is a strong correlation (r=0.87) between the number of scans between question and content to reading behavior observed. Participants who scanned between question and the reference text frequently also exhibited a larger amount of

reading fixation transitions comparative to those that had fewer scans. This indicates that participants who scanned between question and text regions did so to read the text as a reference. From this observation we can conclude that region scans can be used as a metric for answer-seeking behavior.

We propose two purposes for measuring answer-seeking behavior. The first is as a feedback tool for instructors about the nature of how students read and answered questions. The second is to provide feedback to instructors about how individual students are performing. We start by elaborating on the first use; feedback about individual questions. The average number of region scans between question and content, the average number of fixation transitions classified as part of reading behavior and the average score for that question are shown in Table I. It can be seen that there is a large range in the average number of region scans for each question. There is also a large range in fixation transitions classified as part of reading behavior for each question. There is a minimum of 3.3 scans and 23.9 reading transitions for question 3 and a maximum of 19.5 scans and 203.3 reading transitions for question 8. From these observations, question 3 was the easiest on average question to answer as fewest region scans and lowest amount of reading was needed to answer the question. Question 8 was the most difficult question to answer on average as the most region scans and the most reading was needed to answer the question.

TABLE I. ANSWER SEEKING BEHAVIOUR AVERAGES PER QUESTION.

| Question number | Region Scans between question and content | | Transitions classified as part of Reading Behavior | | Average score for question | |
|---|---|---|---|---|---|---|
| | *Mean* | *STD* | *Mean* | *STD* | *Mean* | *STD* |
| 1 | 11.1 | 9.5 | 85.9 | 79.4 | 0.73 | 0.46 |
| 2 | 12.2 | 9.2 | 115.6 | 118.6 | 0.90 | 0.21 |
| 3 | 3.3 | 3.4 | 23.9 | 36.7 | 1.00 | 0.00 |
| 4 | 5.9 | 2.6 | 65.5 | 52.8 | 1.00 | 0.00 |
| 5 | 4.6 | 4.6 | 36.9 | 40.3 | 0.80 | 0.41 |
| 6 | 4.3 | 3.2 | 32 | 20.8 | 1.00 | 0.00 |
| 7 | 10.6 | 8.0 | 58.6 | 47.7 | 0.87 | 0.35 |
| 8 | 19.5 | 13.6 | 203.3 | 173.5 | 0.90 | 0.21 |
| 9 | 6.1 | 5.0 | 73.5 | 90.9 | 0.73 | 0.46 |
| 10 | 11.2 | 8.1 | 77.5 | 55.0 | 0.97 | 0.13 |
| 11 | 17.3 | 12.9 | 148.1 | 107.9 | 0.73 | 0.46 |
| 12 | 17.3 | 11.5 | 146.7 | 120.6 | 0.97 | 0.13 |
| 13 | 4.6 | 4.1 | 57.3 | 60.6 | 0.80 | 0.41 |
| 14 | 7.9 | 5.6 | 65.9 | 80.4 | 1.00 | 0.00 |
| 15 | 5.8 | 7.7 | 64.8 | 100.4 | 0.67 | 0.49 |
| 16 | 7.1 | 4.2 | 61.2 | 46.0 | 0.97 | 0.13 |
| 17 | 5.5 | 4.2 | 33.7 | 32.7 | 0.93 | 0.26 |
| 18 | 9.3 | 5.3 | 71.7 | 42.8 | 0.93 | 0.18 |

On average, participants displayed more region scans and transitions classified as part of reading behavior observed for some questions; that is, more answer-seeking recorded. This indicates that some questions are harder than others to answer. This difficulty could be for several reasons such as ambiguity in the question or technical difficulty of the question. There is no correlation between the number of region scans between questions and content or the amount of reading transitions observed to the score obtained for the question (r=-0.1, for both measures). Therefore, performance on the question is not an accurate measure of how difficult the participants found the questions. Instead the answer-seeking behavior is a measure of how hard the participants found the questions as well as how much interest they paid to the question. We propose the use of answer-seeking behavior to describe how difficult a question is to answer.

The large standard deviations shown in Table I show that there is a large variation in the observed answer-seeking behavior. This is expected, as there is a large variation in eye movement behavior observed between individuals [1]. Furthermore, we are only considering average performance on questions, as we would expect that some individuals would find questions easier to answer than others. This leads us to the discussion of using answer-seeking behavior to quantify individual student performance.

Next we investigate the use of answer-seeking to establish participant behavior. We have established in previous work that answer-seeking behavior is an indicator of a participant's confidence in answering the questions. The average number of region scans between question and content, the average number of fixation transitions classified as part of reading behavior, and the total score for that participant are shown in Table II. Note that in Table II the participants are listed in ascending order of average number of scans between questions and content. This ranking of participants' shows the extent of the variance of answer-seeking behavior each participant exhibits. Once again we can use this information to extrapolate how difficult the individual participant found the tutorial and quiz. The 15$^{th}$ participant showed quite a high amount of answer seeking behavior whilst the 1$^{st}$ participant showed about a seventh of the region scanning and transitions classified as reading to the 15$^{th}$ participant. There is a small negative correlation between the average number of region scans and the participants' total score (r=-0.34). This indicates that the participants who displayed less answer-seeking behavior were not necessarily correct and may be over confident with their answers.

TABLE II. AVERAGE ANSWER SEEKING BEHAVIOUR PER PARTICIPANT.

| Participant | Region Scans between question and content | | Transitions classified as part of Reading Behavior | | Total Score |
|---|---|---|---|---|---|
| | *Mean* | *STD* | *Mean* | *STD* | |
| 1 | 2.9 | 3.2 | 24.9 | 34.9 | 17 |
| 2 | 3.2 | 3.6 | 23.7 | 21.9 | 18 |
| 3 | 5.0 | 3.4 | 67.4 | 68.3 | 17 |
| 4 | 5.4 | 4.7 | 46.6 | 55 | 17 |
| 5 | 6.4 | 4.7 | 40.3 | 38.9 | 16.5 |
| 6 | 6.8 | 5.8 | 47.1 | 43.1 | 13 |
| 7 | 6.9 | 5.9 | 39.4 | 43.4 | 17 |
| 8 | 7.8 | 6.8 | 95.8 | 108.1 | 14.5 |
| 9 | 8.6 | 5.6 | 60.2 | 51 | 18 |
| 10 | 10.6 | 11.4 | 89.7 | 103.6 | 14.5 |
| 11 | 11.4 | 6.9 | 136.2 | 57.9 | 15 |
| 12 | 12.3 | 8 | 56.7 | 51.9 | 13.5 |
| 13 | 12.3 | 9.7 | 143.7 | 163.8 | 16 |
| 14 | 14.8 | 8.6 | 139.4 | 95.1 | 15.5 |
| 15 | 22.1 | 14.3 | 173.9 | 129.6 | 16 |

We propose the use of answer-seeking behavior as an implicit measure of how difficult a participant finds the tutorial and quiz.

Once again there is high variation in the observations as shown by the standard deviations. This is a reflection of the differing complexity of the 18 questions as already discussed and shown in Table I. The standard deviations for each participant can be used to evaluate how consistently difficult that participant found the questions. For example, a low standard deviation indicates low variability and therefore that the participant consistently showed similar answer-seeking behavior. This result indicates that the participant found each question to be similar in complexity. If it is an objective for instructors to construct questions of similar or differing complexity than this information can be used in support of the previous information about question complexity (shown in Table I).

## V. Predicting Reading Comprehension

We now investigate how answer-seeking behavior and eye movement measures can be used to predict participant comprehension. Both the participants' subjective comprehension and quiz score are used as comprehension metrics. The technique used to perform the prediction is a feed-forward backpropagation neural network. Mean square error (MSE) was used to evaluate the performance of the network output. The reported MSE values are the average of 50 runs of the network. Each time 10 participants are selected randomly as the training set and 5 participants randomly selected as the testing set. We used the average answer-seeking behavior and eye movement measures as inputs to the network. The eye movements measures used are: number of fixations, maximum fixation duration, average fixation duration, total fixation duration, number of regressions and regression ratio. There are a total of 16 input neurons to the network and 16 hidden neurons. The average MSE for the training set for predicting participant quiz score is 0.74 and the average MSE for the testing set is 4.3. The average MSE for the training set for predicting participant subjective comprehension score is 0.85 and the average MSE for the testing set is 3.8.

These results show that both subjective and explicit comprehension scores can be predicted using answer-seeking behavior and eye movement measures. The implication of predicting subjective comprehension based on eye movements is that feedback can be given to both instructors and students about their relative understanding of the content and questions. To elaborate on this, if a student is exhibiting eye movement behavior that implies he or she subjectively believes they understand the content however their explicit comprehension scores show otherwise, retargeting of the educational material to the student can be made before giving them a test result. The upshot of this is that the student is not unknowingly demoralized by a low result, which could affect subsequent learning behavior. Instead students would be encouraged to learn the material in a more positive way rather than providing negative feedback.

## VI. Using Answer-Seeking Behavior for Feedback

We have defined an indicator of reading intensity by number of fixations, number of regressions, and total fixation time. We have also shown that reading intensity can be used as a supporting indicator for reading comprehension level as more reading intensity for the first read through of the text was seen to correlate to lower reading intensity of the second read through of the text. As stated prior, reading intensity can be affected by many different factors such as how interested the reader is in the content, and how hard the reader finds the content to understand. These factors also contribute to reading comprehension, as readers are more likely to want to learn something they are interested in and are less likely to understand something that is complicated to read. In this case, eye tracking data from students reading learning material can be used to calculate reading intensity of learning material. The reading intensity can then be used along with answer-seeking indicators to extrapolate information about learning content. For example, if a section of text is presented to students and the reading intensity is observed to be low yet later when asked questions based on content, students exhibited high answer-seeking behavior then the instructor can investigate whether the content is interesting enough to keep the students' attention. Another example is if high reading intensity is observed followed by high answer-seeking this could be a signal that the content is complex.

We have established a method of defining answer-seeking behavior by recording the large scans between questions and content combined with the amount of reading that is performed in the question and content areas. We propose the use of this measure as an indicator to the instructor of question difficulty as well as the participant's implicit difficulty in completing the quiz. We will now establish the benefits of such information.

There is a range in answer-seeking behavior seen for each of the questions. This shows that some questions were harder to answer then others. The use of answer-seeking behavior as a measure of question difficulty can be used as a feedback system to an instructor. If such information is provided to the instructor then the instructor could gauge how difficult questions are. This difficulty could be due to factors such as the technical nature of the material, and ambiguity in the material. Conversely, the instructor could see that the question is too easy and change it to be more challenging. This information could also be used to weight questions so that more difficult questions are weighted higher than those that are less difficult.

Furthermore, there is a range of answer-seeking behavior seen among the participants. Some found the quiz more challenging than others. It is beneficial for learning if all students are challenged equally. Under-challenged students may get bored and lose interest in the material whilst over-challenged students may become anxious and disheartened by the material. In either case there is a negative impact on the learning process. Using answer-seeking behavior as an implicit measure for a student's confidence in the material can provide the framework for an adaptive online learning environment. Such an environment can use input from the

eyes to measure the answer-seeking behavior and alter the learning material and questions in response to the student's behavior. That is, if a student is found to not be having difficulty completing a quiz then the material can be altered to be more advanced and technical. Conversely, if a student is having difficulty then the material can be altered to be less technical and more basic.

## VII. Conclusion and Further Work

In this study we have investigated answer-seeking behavior and methods for detecting reading comprehension for a tutorial and quiz. This work builds on previous work on the effect of topic familiarity on eye movements during reading. In this research we identified and measured answer-seeking behavior. We found that participants with higher numbers of fixations and regressions, as well as longer total fixation time during reading of the first presentation of the content had lower respective values for reading the second presentation of the content. These eye movement measures can be used as an indicator of reading intensity, which we propose is related to comprehension.

We have shown that more answer-seeking behavior is observed for those participants who did not answer the questions correctly. Answer-seeking behavior is related to the participant's confidence in answering the question. We have shown that this confidence is also related to whether he or she answered the question correctly or not. Answer-seeking behavior is an indicator of comprehension.

Furthermore, we have proposed the use of answer-seeking behavior to describe how difficult a question is to answer and as an implicit measure of how difficult a participant finds the tutorial and quiz. The eventual goal is to create a tool that will provide feedback to instructors about implicit behavior of students performing a reading task through an online learning environment. For example, if the instructor receives feedback that multiple students are failing to understand specific parts of the text then the instructor can dedicate more time explaining these concepts during face to face teaching time, or could re-word the content to make it easier to understand. Furthermore, the instructor can be given feedback about how students are reading questions and be able to deduce if questions are appropriately worded or are ambiguous and hence causing low scores or confusion. Finally, the information about reading behavior can also be used to dynamically alter tutorial content to personalize the learning experience where students familiar with or excelling at specific content can be given more advanced content to read compared to students that are not familiar with the content or finding it harder to understand.

The outcomes of this current study are aimed to predict a student's reading behavior and understanding of text. We have shown that by using a feed-forward backpropagation neural network both subjective and explicit comprehension scores can be predicted based on answer-seeking metrics and eye movement measures.

This data will be analyzed further to assess how eye movements can be analyzed to assess a participant's understanding of the content and the questions. This is the first step in identifying differences in comprehension processes and formulating ways for how to evaluate them using eye movements as an implicit comprehension measure.

The results from this study will be used as the foundation for building a prototype adaptive learning environment. The learning environment will use eye tracking as the input for determining learning rates and behavior during reading so that learning can be personalized for individual students. This is on-going research, which is currently being explored.

## VIII. References


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

[2] W. Kintsch, and K.A. Rawson, Comprehension, in The Science of Reading: A Handbook, M.J. Snowling and C. Hulme, Editors, 2005. Blackwell Publishing.

[3] G. Underwood, and V. Batt, Reading and Understanding, 1996. Massachusetts, USA: Blackwell Publishers.

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

[5] L. Frazier, and K. Rayner, " Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences." in Cognitive Psychology, vol. 14(2), pp.178-210, 1982.

[6] S.T. Iqbal, and B.P. Bailey, " Using Eye Gaze Patterns to Identify User Tasks." In Proc. The Grace Hopper Celebration of Women in Computing, 2004.

[7] J. Salojarvi, K. Puolamaki, J. Simola, L. Kovanen, I. Kojo, and S. Kaski, "Inferring relevance from eye movements: Feature extraction" in Publications in Computer and Information Science, 2005.

[8] C.S. Campbell, and P.P Maglio., " A robust algorithm for reading detection." in Proc of the 2001 workshop on Perceptive user interfaces. 2001. ACM.

[9] C.J. Gustavsson, Real Time Classification of Reading in Gaze Data (Master's Thesis), 2010, School of Computer Science and Engineering. Royal Institute of Technology. Stockholm, Sweden.

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

[11] T. Vo, B.S.U. Mendis, and T.D. Gedeon, "Gaze Patterns and Reading Comprehension" in Proc. Of the 17th international conference on Neural information processing: Models and applications, vol. II, pp.124-131, 2010.

[12] R.J. Jacob, and K.S. Karn, "Eye tracking in human-computer interaction and usability research: Ready to deliver the promises." in Mind, vol. 2(3): p. 4, 2003.

[13] G. Buscher, A. Dengel, R. Biedert, and L. Van Elst, "Attentive Documents: Eye Tracking as Implicit Feedback for Information Retrieval and Beyond." in ACM Transactions on Interactive Intelligent Systems, vol. 1(2), p. Article 9, 2012.

[14] A. Hyrskykari, P. Majaranta, A. Aaltonen, and K.-J. Räihä, "Design issues of iDICT: a gaze-assisted translation aid." in Proc. of the 2000 symposium on Eye tracking research & applications, pp. 9-14, 2000.

[15] J.L. Sibert, M. Gokturk, and R.A. Lavine, "The Reading Assistant: Eye Gaze Triggered Auditory Prompting for Reading Remediation." in Proc. of the 13th annual ACM symposium on User interface software and technology, pp. 101-107, 2000.

[16] L. Copeland and T. Gedeon, "The Effect of Subject Familiarity on Comprehension and Eye Movements during Reading", in Proc. Of the 25th Australian Computer-Human Interaction Conference, 2013

[17] J.M. Fletcher, "Measuring Reading Comprehension" in Scientific Studies of Reading, vol. 10, pp. 323-330, 2006.