IEEE TRANSACTIONS ON
EMERGING TOPICS
IN COMPUTING

Received 1 March 2015; revised 27 September 2015 and 1 November 2015; accepted 2 November 2015.
Date of publication 16 December 2015; date of current version 3 March 2017.

Digital Object Identifier 10.1109/TETC.2015.2499257

Tutorials in eLearning—How Presentation Affects Outcomes

LEANA D. COPELAND AND TOM D. GEDEON, (Senior Member, IEEE)

Australian National University, Canberra, ACT 2601, Australia

CORRESPONDING AUTHOR: L. D. COPELAND (leana.copeland@anu.edu.au)

ABSTRACT The presentation of learning materials affects how we learn. In this paper, we use eye tracking to investigate how different sequences of text and test questions affect performance outcomes, eye movements, and reading behavior for first (L1) English language and second (L2) English language readers. We show that different presentation sequences induce different performance outcomes, eye movements, and reading behavior. The sequence can affect how a participant reads the text as well as their perceptions of how well they understood what they read. For instance, if questions and text are not shown together, this improves participants' ability to accurately perceive their comprehension and promotes thorough reading. Alternatively, showing questions before the text promotes skimming behavior. Importantly, the presentation sequence affects both L1 and L2 readers in the same way. We observe that L2 reader take longer to read text but have the same comprehension levels as L1 readers; this difference comes primarily from longer fixation durations. The results from this paper can be used to design learning materials in eLearning environments to influence how students interact with the learning environment as well as how they learn. The purpose of this investigation is to make informative decisions about designing adaptive eLearning environments.

INDEX TERMS Human-centered computing, Web-based interaction, education, computer uses in education.

I. INTRODUCTION

Online learning extends teaching and learning from the classroom to a wide and varied audience that has different needs, backgrounds, and motivations. In tertiary education online learning materials are becoming increasingly ubiquitous. This is due in part to increased availability of computer technologies. Universities frequently offer online and/or off-campus degrees where students may have little or no face-to-face interaction with their instructors or other students. Even for university courses that deliver educational material traditionally, absenteeism from lectures is more prevalent and has been shown to negatively affect learning [1], [2]. Additionally, access to texts that are not necessarily written in a reader's native language is now easy and common. There are known differences in the reading behaviors of first (L1) and second (L2) English language readers [3]. For these reasons is a growing importance in designing effective eLearning materials that take into consideration the differences between L1 and L2 readers.

The way in which learning materials are presented to students can have great bearing on the outcomes of comprehension. It has been established that the presentation of images with text increases comprehension [4]. Moreover, pretesting with multiple choice questions improves subsequent learning of materials [5]. We explore further how presentation of learning materials affects reading and learning behavior. We investigate the effects that test questions have on learning by presenting questions and text in different sequences. Furthermore, we investigate if the effects of sequence are different for L1 and L2 readers since there is a growing diversity in the audiences of eLearning courses.

To investigate this, a user study was conducted to compare four different sequences of text and test questions. In the user study, participants' eye gaze was recorded, using eye tracking technology, as they read text and answered questions. Eye tracking has been shown to be an effective way of analyzing various human behaviors, particularly reading (see review by [6]). Eye movements are unique in reading and can reveal when readers encounter difficulties in reading [7] as well as text difficulty and comprehension [8]. Eye tracking is especially useful at analyzing the implicit differences between different types of readers, such as L1 and L2 readers [3].

We hypothesize that the presentation sequences of the text and test questions will: 1) affect L1 and L2 readers in the same way, however, there will be differences between the two groups; 2) have an effect on participants' performance, in terms of time and quiz score, and perceived understanding of the text; and 3) cause differences in eye movements and induce different reading behavior.

This paper is organized into the following sections: background information; user study methodology; results and analysis; discussion and recommendations; finishing with conclusions and further work.

II. BACKGROUND

A. READING IN A DIGITAL WORLD AND ONLINE LEARNING

The rise of the Internet and expansion of the World Wide Web has seen an increase in reading in many countries [9]. This increase coincides with the proliferation of mobile technology such as smart phones and tablets. The Internet is now available almost anywhere and anytime if you have a smart device. The debate about the effects of such digitization and rapid access to vast quantities of information ranges from ergonomics [10], [11], effect on memory [12], reading comprehension and effects on learning [13]–[16].

1) ELECTRONIC TEXT (eTEXT)

Electronic text (eText) is the general term for digital presentation and storage of text that can be read via a digital device, such as a computer, laptop, tablet, smart phone, or eReader. The digitization of text has spawned a great deal of research into the effects this has on the reading process. Initially, much research went into comparing reading digital to paper based texts [10], [17]. We now give a brief overview of differences that have been found in the context of educational materials.

Hypertext is a prominent form of eText as it is the primary form delivery of information on the web. A hypertext document enables the reader to navigate in a non-liner way via links to other resources or pieces of text. The resulting structure of a hypertext document can be complex. The consensus is that hypertext structure negatively impacts the reading process due to increased cognitive demand needed for decision making and visual processing [13].

Other eTexts include PDFs (portable document format) or eBooks (electronic books), which are linear and therefore closer to print media. Whilst comprehension of text in print form is significantly higher compared with PDF form [15], students who purchase electronic textbooks perform no differently in a university course [16].

Paper offers advantages over digital text such as supporting annotation, quick and easy navigation, as well as control of spatial layout [18]. However, eText does itself have advantages over paper, such as increased accessibility, easy storage and retrieval, ubiquity, and flexibility. Flexibility refers to the ability to dynamically change the way text is read. Changes can be simple, for example changing of font size, color, or typeface. Changes can also be complex, such as verbalizations of the text, embedded definitions, and links to background information [19]. The ability to dynamically change eText presents the opportunity to make transformations to promote learning and comprehension.

Many studies have examined navigation through eTexts as it is often non-trivial [11]. Studies have investigated navigation in eBooks [20], academic articles [17], and periodicals [21], as well as the impact of screen size on document triage [21]. Also, the impacts of highlighting, hyperlinks, fonts, distractions (such as alerts), as well as embedded videos and sounds have been extensively investigated. The insight gained from these studies is beneficial in designing online reading materials. Inappropriate highlighting of words negatively affects reading comprehension whereas appropriate highlighting enhances comprehension [22]. The effects of font and font size used in eText have been investigated, where the focus has been on comparing serif and san-serif fonts [23]–[25]. Smaller font sizes tend to induce slower reading speeds [23], [24]. This was found to result from increased fixation duration [24].

2) ONLINE eLEARNING

Online learning material is accessible virtually anywhere and anytime to a wide and varied audience. Students are increasingly skipping lectures in favor of accessing digital copies or recordings of the lectures. Indeed, the availability of lecture webcasts and PowerPoint slides negatively impacts student attendance [26], whilst missing face-to-face tuition can have a negative effect on learning [1], [2].

The advent of massive open online courses (MOOCs) has also increased importance in designing effective eLearning materials. MOOCs have become popular in the past couple of years. The goal of MOOCs is to provide free or low cost but quality education that is available to anyone who wishes to take part. There are now many examples of reputable websites that offer MOOCs, such as Udacity, Coursera, edX, and Khan Academy. Whilst on one hand MOOCs do achieve the goal of making educational resources available to people who would not have access to them otherwise, they suffer from extremely low completion rates. The completion rate of edX's first MOOC is below 5% [27]. One of the problems identified with MOOCs is that they are indeed massive and open and so are easy to get lost in. Hence, there is a growing importance to design more effective eLearning materials.

B. EYE MOVEMENTS AND READING COMPREHENSION

During the process of reading the eye moves in well-studied ways that can be broadly characterized as fixations and saccades. A fixation is where the eye remains relatively still to take in visual information and a saccade is a rapid movement that transports the eye to another fixation. This is because only a small part of the retina actually sees in fine detail. This part is called the fovea and is only 0.2mm in diameter. Around the point of fixation, visual acuity extends about 2° [6]. Humans see very little in detail at any fixation so the eye must move around rapidly so that it can compose a more detailed view of the environment.

When reading English fixation duration ranges anywhere between 60-500 milliseconds and is generally about 250 milliseconds [28]. Saccadic movement is between 1 and 15 characters with an average of 7-9 characters. The majority of saccades are to transport the eye forward in the text when reading English; however, a proficient reader exhibits backward saccades, regressions, to previously read words or lines about 10-15% of the time. Long regressions occur due to comprehension difficulties [7].

Eye movements can be used to understand the ongoing cognitive processes that occur during reading [28]. Comprehension of text can have significant effects on eye movements [8]. A number of studies have shown there are many variables based upon comprehension functions that can influence eye movements during reading. These variables include: semantic relationships, anaphora and co-reference, lexical ambiguity, phonological ambiguity, discourse factors, stylistic conventions, and syntactic disambiguation (see review by [6]). These variables have different effects on eye movement, causing them to deviate from the default reading process. For example, syntactically ambiguous sentences induce regressions to resolve the comprehension problems [7]. Eye movements have been shown to reflect global text difficulty as well as inconsistencies within text [8]. More difficult text causes more fixations, more regressions, and longer fixation duration time. Eye movement has also been shown to indicate reading comprehension and reading skill [29].

C. IMPROVING eLEARNING: PROVIDING ADAPTIVITY

The increased use and diversity in audiences of eLearning environments means that there is a growing importance to design more effective eLearning materials, yet eLearning for the most part is one size fits all. One solution is to make eLearning adaptive to each student. Adaptive eLearning is not new and has already started to show great promise in improving education. Adaptive tutorials have been harnessed to decrease failure rates in early year engineering subjects and drastically increase student enrollment and satisfaction [30].

Adaptivity is achieved through different means including students' measured skill level [31], questionnaires [32], prediction of learner style [33]–[35]. Whilst traditionally the learning style is determined via questionnaire [32] there is progressively more research showing that measures of the students behavior and biometric technology can predict learning style [34], [35]. Mouse movement patterns have been shown to have a high correlation and global/sequential learning style [35]. Eye tracking has also been shown to be a potential way of identifying visual/verbal learner style [34]. Eye movements in areas of interest on the page were related to measures of learner style in that investigation. Similar uses of eye tracking have been used to investigate learning behaviors between novice and advanced students when learning SQL [36]. Advanced students looked at the database schema more that novices. Studies such as this are useful for identifying such difference in order to provide more help for novice students.

The concept of adaptive eLearning also extends to mobile learning. The MAPLE framework uses a combination of eye tracking and accelerometer data to determine learner style in both mobile and online environments [37]. Generic Responsive Adaptive Personalized Learning Environment (GRAPPLE) project [38] is another adaptive learning environment through adaptive guidance and personalized learning content. The authors of GRAPPLE show how they can integrate their system with currently used LMSs such as Moodle. Other frameworks take into account students with learning problems such as dyslexia [39]. The Dyslexia Adaptive eLearning (DAEL) framework is designed to be tailor the learning material according the dyslexia type [39].

Adaption is not only provided via detection of learning style. Eye tracking can be used to detect many facets of human behavior. Eye gaze patterns have be used to detect what kind of task the participant is performing [40] or whether a person is reading or not [41] as well as if they are reading or just skimming [42]. More recently, eye gaze has been used to investigate parts of text that readers are failing to comprehend text [43]. Eye tracking can also be used to analyze how multiple choice questions are answered [44], [45] and to predict student performance of physics concepts when presented as text or images [46].

Eye tracking has been used in multiple ways to provide adaptivity to eLearning. A classic example of the use of eye tracking in eLearning is AdeLE (Adaptive e-Learning with Eye-Tracking). The AdeLE project sets out a structure for how an adaptive eLearning environment could be constructed using eye tracking data such as blink rate and how open the eye lid is [47].

One aspect that has about learning that is frequently investigated is engagement. Eye tracking has been used to identify aspects of a student's emotional state, such as stress and arousal, and adapt the material based on the identified state. An example of this is e5Learning (enhanced exploitation of eyes for effective eLearning) which uses eye gaze metrics such as fixation statistics and pupil diameter to identify the students emotional state [48]. Gaze Tutor uses eye gaze to determine the user's level of stimulation to alter the environment to stimulate the user [49]. An interesting approach to identifying students engagement comes from the use of a type-2 fuzzy logic based system [50]. This novel method gauges degree of engagement to adapt the learning environment. Results show that using the system to adapt material there is significant improvement in average scores compared to other methods of adaption and no adaption.

Eye tracking is also used to analyze reading in eLearning environments. One example is iDict, a reading aid designed to help readers of a foreign language that uses eye gaze to predict when a reader is having comprehension difficulties [51]. If the user hesitates whilst reading a word then a translation of the word is provided along with a dictionary meaning. Similarly, the Reading Assistant [52] uses eye gaze to predict failure to recognize a word. The Reading Assistant then provides an auditory pronunciation of the word to aid reading. Eye movements have been used to monitor reading behavior in combination with measuring pupil size as a means of gauging mental workload [53].

III. METHODOLOGY

A. DESIGN

Our study used a between-subjects design where participants were shown one of four presentation sequences of text and questions. The independent factors of the experiments are the presentation type and English as a first (L1) or second language (L2).

B. MATERIALS & PROCEDURE

The user study conducted involved tracking participants' eye gaze as they read a text and answered questions. The text and questions are a tutorial from a first year Computer Science course run at the Australian National University. The tutorial and quiz is composed of 9 screens of textual content, each covering a specific area about the main topic of the tutorial (Web Search). Each text is 400 words long and has an average Flesch Kincaid Grade readability level of 11.5. This indicates that participants need a 12th grade education level. As the slides are targeted at a first year university students this is a suitable readability level. For each screen there were two comprehension questions; one multiple-choice and the other cloze (fill-in-the-blanks). These two types of questions were used because they can be used to assess different forms of comprehension [54].

The tutorial and quiz was presented to participants in four formats to measure the effect of presentation on participants' eye gaze and answering behavior. These presentation formats are described below:

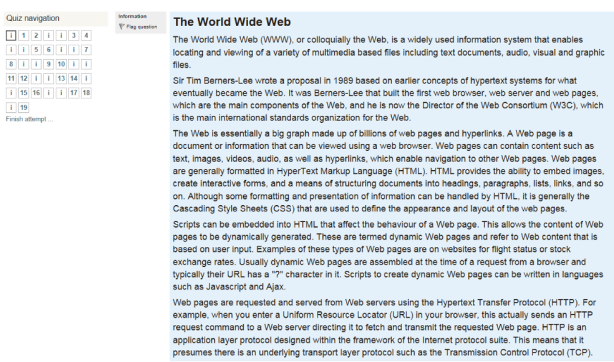

FIGURE 1. Example of tutorial text only slide.

1) **Format A ($\mathbf{T \rightarrow T/Q}$).** The text slide (shown in Fig. 1.) is first shown to participants followed by a slide with both questions and text (shown in Fig. 2.). Since there are 9 topics there are 18 slides in total. Participants are required to read the text before being able to read the questions relating to it. When participants view the questions they can either answer straight away or search the text to look for the answers, i.e. answer-seeking behavior [55], [56].

FIGURE 2. Example of tutorial text and question slide.

2) **Format B ($\mathbf{T/Q}$).** The questions and text slide (shown in Fig. 2.) is shown to participants. Since there are 9 topics there are 9 slides in total. Participants are no longer required to read the text before they see the questions. Our question is will participants read the text completely or will answer seeking behavior be observed? Additionally, is there a difference in quiz performance when participants can immediately answer the questions without reading the text?

3) **Format C ($\mathbf{T \rightarrow Q}$).** The test slide (shown in Fig. 1.) is first shown to participants followed by the questions only slide (shown in Fig. 3.). Since there are 9 topics there are 18 slides in total. This format is the control presentation method. In this format the reference text is removed from the questions slide so the participants are forced to answer the questions from understanding and memory. We expect that the worst comprehension scores will be observed for this format. Format C is the most commonly used in on-line quizzes.

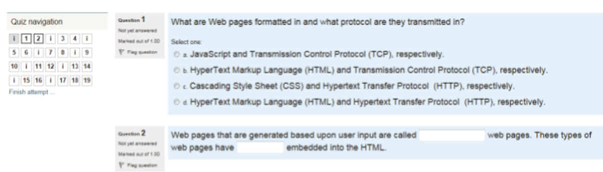

FIGURE 3. Example of the questions only slide.

4) **Format D ($\mathbf{Q \rightarrow T \rightarrow Q}$).** The last presentation format consists of displaying a slide with only the questions (shown in Fig. 3.) followed by the text slide (shown in Fig. 1.) and then again presenting them with the

questions only slide. Since there are 9 topics there are 27 slides in total. The hypothesis is that participants will read the text differently than for formats A and C.

C. PARTICIPANTS

The study included 60 participants who were divided equally into the four groups, each of which was shown one of the presentation formats. The breakdown of participants into groups is as follows:

1) **Format A.** 15 participants (6 female, 9 male) with an average age of 22.3 years (standard deviation 4.1 years, range 17-31 years). English was not the first language for 4 of the participants.
2) **Format B.** 15 participants (6 female, 9 male) with an average age of 22.7 years (standard deviation 6.0 years, range 18-41 years). English was not the first language for 4 of the participants.
3) **Format C.** 15 participants (5 female, 10 male) with an average age of 23.5 years (standard deviation 5.3 years, range 18-37 years). English was not the first language for 6 of the participants.
4) **Format D.** 15 participants (7 female, 8 male) with an average age of 22.2 years (standard deviation 3.3 years, range 17-28 years). English was not the first language for 5 of the participants.

D. EXPERIMENT SETUP

The quiz was accessible via a Moodle variant, the online learning environment used at our university. 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 60 milliseconds and a threshold of 5 pixels.

E. DATA PRE-PROCESSING

The raw eye gaze data consists of x,y-coordinates recorded at equal time samples (60Hz). Fixation and saccade identification was performed on the eye gaze data. From this data many other eye movement measures are derived. The measures used in this analysis are:

1) **Number of fixations:** The sum of fixations recorded for each page. The number of fixations can be affected by reading behavior, text difficulty, and reading skill [6].
2) **Maximum fixation duration (seconds):** The maximum duration of the longest fixation recorded for each page. Longer fixations can be an indicator of difficulties in processing particular words or due to linguistic and comprehension difficulties [6].
3) **Average fixation duration (seconds):** The sum of the duration of all fixations divided by the number of fixations. This measure has been used to predict reading comprehension [57].
4) **Total fixation duration (seconds):** The sum of all fixations. This measure is useful in global text processing analysis [58].
5) **Number of regressions:** The number of regressions. When reading more difficult text more regressions are observed [8].
6) **Reading analysis:** Using our combination of two reading detection algorithms [41], [42] the percentage of saccades classified as reading (read ratio), skimming (skim ratio), and scanning/searching (scan ratio) is calculated.

Participants' quiz outcomes are measured to assess how well they performed under different conditions. The measures of participants' performance are:

1) **Subjective comprehension:** a self-rated measure between 0 and 10, where 10 is comprehensive understanding of the material.
2) **Comprehension question scores:** the multiple-choice questions are graded as 0 (incorrect) or 1 (correct) and the cloze questions are scores as 0 (incorrect), 0.5 (one word was correct) or 1 (correct). The maximum total score for the quiz is 18.
3) **Time taken (minutes):** the total time it took each participant to complete the tutorial and quiz is recorded.

IV. RESULT & ANALYSIS

This section contains a statistical analysis of participants' performance outcomes and eye movements under each of the experimental conditions. This section is structured to address the two hypotheses that presentation effects measured performance and eye movements. In both of sections the differences between L1 and L2 readers is addressed.

A. DOES FORMAT AFFECT PERFORMANCE?

The question of whether format affects performance incorporates two hypotheses that will be explored in this subsection. Firstly, we hypothesize that the different presentation formats will affect participants' scores, time taken to complete, and perceived understanding, and these effects will be the same for both L1 and L2 readers. Secondly, only time taken to complete will be different between the L1 and L2 readers.

The mean and standard deviations for time taken (minutes) to complete the tutorial and quiz is shown in Fig. 4; the quiz grade is shown in Fig. 5; and the participants' subjective understanding is shown in Fig 6.

To address the hypotheses a MANOVA is used to determine if there are any statistical differences between the formats and L1/L2 readers. The correlations between the dependent variables are within the acceptable limits for MANOVA outcomes, i.e. the correlations lie between $r = -0.4$ and $r = 0.9$. To test for normality in the dependent variables the Shapiro-Wilk Test is used, as it is more appropriate for small sample sizes. The quiz scores are normally distributed for all formats. The times taken are

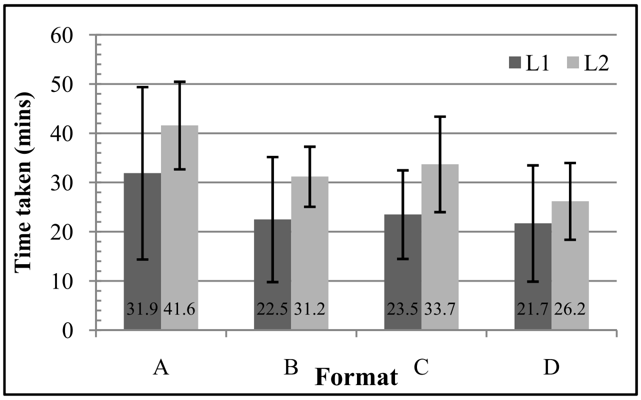

FIGURE 4. Means and standard deviations of time taken to complete the tutorial for each format. (A: $T \rightarrow T/Q$; B: T/Q; C: $T \rightarrow Q$; D: $Q \rightarrow T \rightarrow Q$).

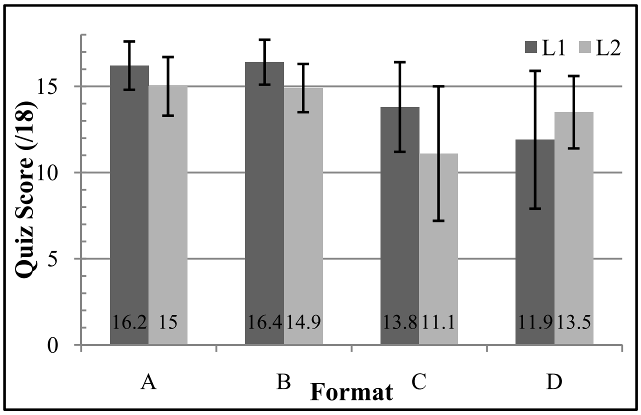

FIGURE 5. Means and standard deviations of quiz scores for each format. (A: T → T/Q; B: T/Q; C: T → Q; D: Q → T → Q).

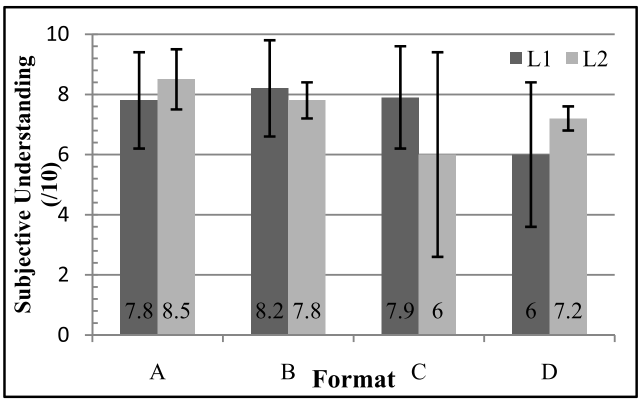

FIGURE 6. Means and standard deviations of subjective understanding scores for each format. (A: $T \rightarrow T/Q$; B: T/Q; C: $T \rightarrow Q$; D: $Q \rightarrow T \rightarrow Q$).

normally distributed for the formats A, B and C, it is just the times for D ($p = 0.026$) that could be a problem. Whilst the subjective scores for B are normally distributed and the scores for C are very close to being normally distributed, the scores A and D could be a problem. Finally, the homogeneity of variance-variance-covariance matrices is satisfied as the Box's M value of 69.73 ($p = 0.165$).

1) EFFECT OF FORMAT ON PERFORMANCE MEASURES

The results indicate format affects performance outcomes for both L1 and L2 readers. This is supported by the results from the MANOVA that show there is a significant difference in performance variables based on the format, $F(9, 121.8) = 4.036$, $p < 0.0005$; Wilk's $\lambda = 0.530$, partial $\eta^2 = 0.191$. There is no significant effect of interaction between the format and reader type, indicating that format effects L1 and L2 readers in the same way.

Since significant results were found, between-subjects ANOVAs are performed. Format has a significant effect on both the quiz grade ($F(3, 52) = 6.078$; $p = 0.001$, partial $\eta^2 = 0.260$) and on time taken ($F(3, 52) = 5.552$; $p = 0.002$, partial $\eta^2 = 0.243$), however not on the subjective comprehension score.

Tukey's HSD tests are used to make pairwise comparison of the formats. There is no significant difference between formats A and B or formats C and D. These two groups correspond to similarities in presentation formats. Formats A and B show the questions with the tutorial text whereas formats C and D do not. Two conclusions can be made from this observation; firstly the lack of difference between formats A and B illustrates that reading the tutorial text before being presented with the questions does not improve comprehension scores. Secondly, when comparing formats C and D, the knowledge of the questions before reading the text also does not improve comprehension scores.

Formats A and B have significantly higher quiz scores than formats C and D, (formats A and C ($p = 0.006$), A and D ($p = 0.003$), B and C ($p = 0.005$), and B and D ($p = 0.002$)). For formats C and D the participants must rely on memory and their understanding of the material. Format A takes significantly longer to complete than formats B ($p = 0.011$) and D ($p = 0.002$). However there is no significant difference between the other formats.

The format has no significant effect on subjective comprehension scores. However, for format C and D there are a strong positive correlation between the quiz scores and the subjective comprehension scores ($r = 0.9$ and $r = 0.82$, respectively). In these formats, participants estimate their comprehension level more accurately compared to other formats. Participants shown formats A and B seem unable to estimate their own comprehension levels ($r = 0.32$ and $r = -0.09$, respectively). An important part of the learning process is awareness of skill [59]. Under-estimation of understanding can lead to students wasting time on material already understood instead of using the time to learn more material. On the other hand, overestimation of understanding will results in students not learning what they need to and not realizing their lack of understanding.

2) L1 VERSUS L2 READERS

The second hypothesis is that the only difference expected between L1 and L2 readers will be in time taken. The MANOVA shows that there is a statistical significant difference in performance variables between L1 and L2 readers, $F(3, 50) = 5.79$, $p = 0.002$, Wilk's $\lambda = 0.742$, partial $\eta^2 = 0.258$.

Once again between-subjects ANOVAs are used to compare the groups for each performance variable. The difference between L1 and L2 readers is significant for time taken ($F(1, 52) = 13.135$; $p = 0.001$, partial $\eta^2 = 0.202$) but has no significant effect on subjective comprehension or quiz score. This confirms our expectations and is analogous to existing research that has shown that although L2 readers take longer to read, they perform no differently to L1 readers in comprehension [3]. We have also found there is no difference in their subjective comprehension.

3) SUMMARY

The interim conclusion made from this analysis is that presentation formats affect students' performance. In concordance with current research it was found that L2 readers took longer to complete the quiz but performed no differently to L1 readers. Additionally, the differences in measures caused by formats are consistent for both L1 and L2 readers. The presentation format can be manipulated in the same way for both L1 and L2 reader to optimize the performance outcomes of students, to increase their understanding.

B. DOES FORMAT AFFECT EYE MOVEMENTS?

The final hypothesis is that presentation format affects eye movements and that the eye movements of L1 and L2 readers will be different. To address these overall hypotheses the two central differences in presentation formats are analyzed separately. That is, first the text when shown without the questions will be analyzed and then the text when shown with the questions.

1) ANALYSIS OF EYE MOVEMENT FROM TEXT PAGES

Two types of behavior are hypothesized for reading the text without the questions, first is that when presented with format C, participants more care reading the text than in formats A or D, and secondly, participants presented with format D will skim the text to find the paragraphs where they believe the answers are located and read only those paragraphs thoroughly. Once again, we hypothesize that L2 readers will take longer to read the text. The final hypothesis is a deeper analysis into the observation that L2 readers have longer read times than L1 readers.

A MANOVA is used to check for statistical significance of eye movement measures between formats and reader type. The correlations between the dependent variables are all within the range of $r = -0.4$ and $r = 0.9$. Additionally the majority of the dependent variables are normally distributed according to the Shapiro-Wilk test for normality for both reader type and format. The total fixation duration time was excluded from the analysis as it did not have a normal distribution. The Levene's test for equality of variances shows that there is homogeneity for all dependent variables (significance > 0.05). Additionally, the Box's M value of 98.1 ($p = 0.025$) is interpreted as non-significant so we can be satisfied that we have homogeneity if variance-variance-covariance matrices. The mean and standard deviations for the eye movement measures are shown in Table 1.

TABLE 1. Comparison of eye movement measures for text only pages (mean ± standard deviation) (A: $T \rightarrow T/Q$; B: T/Q; C: $T \rightarrow Q$; D: $Q \rightarrow T \rightarrow Q$).

Format	A		C		D	
Type	L1	L2	L1	L2	L1	L2
Num. Fixations	241±21	311±35	245±23	351±28	178±22	221±31
Max fixation dur (s)	1.1±0.2	2.1±0.3	1.3±0.2	1.6±0.2	1.0±0.2	1.9±0.3
Ave fixation dur (s)	0.17±0.02	0.25±0.03	0.21±0.02	0.23±0.02	0.17±0.02	0.26±0.03
Num. regressions	74±7	83±11	75±7	106±9	66±7	66±10
Read ratio (%)	65±5	73±8	70±5	72±7	48±5	72±7

There is a significant difference in eye movement measures based on format, $F(10, 72) = 3.043$, $p = 0.003$; Wilk's $\lambda = 0.486$, partial $\eta^2 = 0.303$, as well as between L1 and L2 readers, $F(5, 35) = 3.623$, $p = 0.010$; Wilk's $\lambda = 0.659$, partial $\eta^2 = 0.341$. There was no significant effect of interaction between the format and reader type, so format affects both L1 and L2 readers the in the same way.

Between-subjects ANOVAs are used to determine how the eye movements differ for the formats and reader types. Format has a significant effect on the number of fixations ($F(2, 39) = 7.262$; $p = 0.002$; partial $\eta^2 = 0.271$), and number of regressions ($F(2, 39) = 4.234$; $p = 0.022$; partial $\eta^2 = 0.178$), but no effect on maximum fixation duration, average fixation duration or the read ratio.

Tukey's HSD tests are used to make pairwise comparisons of the formats. There is significant difference between number of fixations for formats A and D ($p = 0.034 < 0.05$) and between formats C and D ($p = 0.002$). There is a significant difference between the number of regressions for formats C and D ($p = 0.030$), but not formats A and D or A and C. There is no significant difference between formats A and C for any of the eye movement measures.

It was hypothesizes that for format C participants would read the text more thoroughly compared to the other formats. However, the statistical analysis shows that there is no difference in eye movements for formats A and C. The hypothesis is partially supported as format D has significantly less fixations, so there is less overall reading of the text compared to formats A and C.

L2 readers have significantly more fixations than L1 readers ($F(1, 39) = 11.395$; $p = 0.002$; partial $\eta^2 = 0.226$) as well as longer maximum fixation duration ($F(1, 39) = 13.840$; $p < 0.001$; partial $\eta^2 = 0.262$) and longer average fixation duration

($F(1, 39) = 11.527$; p = 0.002; partial $\eta^2 = 0.228$). L2 readers also have significantly higher read ratios for each format compared to L1 readers ($F(1, 39) = 4.951$; p = 0.032; partial $\eta^2 = 0.113$). This outcome agrees with the observation that L2 readers have longer read times than L1 readers. The analysis of eye gaze shows that this is due to higher numbers of fixations for longer durations.

2) ANALYSIS OF EYE MOVEMENT FROM QUESTIONS AND TEXT PAGES

Format A consists of two presentations of the text, first on its own and then with the questions. The hypothesis is that reading text first in format A will help participants answer the questions and hence will need less reference to the text compared to format B. The mean and standard deviations of eye movement measures for formats A and B are shown in Table 2.

TABLE 2. Comparison of eye movement measures for questions and text pages (mean ± standard deviation) (A: *T* → *T/Q*; B: *T/Q*; C: *T* → Q; D: Q → *T* → Q).

Format Type	A L1	A L2	B L1	B L2
Num. Fixations	225±37	246±61	350±37	429±61
Max fixa-tion dur (s)	0.97±0.13	1.65±0.21	1.31±0.13	1.85±0.21
Total fixa-tion dur(s)	38.1±8.3	53.8±13.7	63.7±8.3	102.9±13.7
Num. regressions	97±14	95±23	149±14	167±23
Read ratio (%)	35±5	50±8	34±5	47±8

A MANOVA is used to test for statistical significance of eye movement measures between formats and reader type. The correlations between the dependent variables are all within the range of $r = -0.4$ and $r = 0.9$. All of the dependent variables are normally distributed according to the Shapiro-Wilk test except average fixation duration which is therefore excluded from the analysis. Levene's test for equality of variances shows that the homogeneity for all dependent variables (significance > 0.05). Box's M value of 45.8 ($p = 0.005$) is interpreted as non-significant so we can be satisfied that there is homogeneity in the variance-variance-covariance matrices.

There is a significant difference in eye movement measures based on format, $F(5, 22) = 3.142$, $p = 0.027$; Wilk's $\lambda = 0.583$, partial $\eta 2 = 0.417$, as well as between L1 and L2 readers, $F(5, 22) = 3.309$, $p = 0.022$; Wilk's $\lambda = 0.571$, partial $\eta^2 = 0.429$. There is no significant effect of interaction between reader type and format.

Between-subjects ANOVAs are used to compare each of the eye movement measures for each of the formats and reader types. Format has a significant effect on number of fixations ($F(1, 26) = 9.279$, $p = 0.005$), total fixation time ($F(1, 26) = 10.924$, $p = 0.003$), and number of regressions ($F(1, 26) = 10.827$, $p = 0.003$), but not on maximum fixation duration or the read ratio. Format B has more observed fixations, longer total fixation time and more regressions than format A. This confirms the hypothesis that less eye movements would be observed for format A.

L2 readers have significantly longer maximum fixation durations ($F(1, 26) = 12.230$, $p = 0.002$) and higher read ratios ($F(1, 26) = 4.350$, $p = 0.040$) compared to L1 readers. Additionally, L2 readers have significantly longer total fixation durations than L1 readers ($F(1, 26) = 5.870$, $p = 0.023$). However now there is no difference between the numbers of fixations of regressions observed for L1 and L2 readers. This is an interesting result as no difference in the number of fixations between L1 and L2 readers indicates that the increase in time taken for L2 readers is due primarily to increase in fixation duration.

C. ANSWER-SEEKING BEHAVIOR

Formats A and B provide participants with the ability to check the content whilst answering the questions. This results in behavior characterized by participants skipping back and forth between the question and the text to find the answer and check that it is correct. This section will begin with a definition of this behavior and then look at how the presentation format can affect this behavior.

1) DEFINING ANSWER-SEEKING BEHAVIOR

In previous research on format A, strong negative correlations between eye movement measures recorded when reading the questions and content page with scores was observed [55], [56]. The hypothesis was that re-reading the questions repeatedly was an indication that there was difficulty answering the question. Participants who do not understand the question or the content well enough to answer the question re-read both the question and the text to make sure they have the right answer. We termed this behavior answer-seeking and proposed it as a measure of participants' confidence in answering questions.

The measures for answer-seeking behavior are the jumps between question and text and the reading that occurs after each jump, shown in Fig. 7. Reading is detected using a combination of reading detection algorithm [41], [42].

Answer-seeking behavior for format A has been thoroughly studied [55], [56]. We now investigate how does the answer-seeking behavior differ between formats A and B? Participants are requested to read the content before moving on to answer the comprehension questions in format A so we postulate that they should have some idea about the answers to the questions as well as knowledge of where to find the answers in the text. The hypothesis is that participants would on average show less answer seeking behavior when presented with format A compared to B. The answer seeking results are show in Fig. 8 and Fig. 9.

A MANOVA is used to test for statistical significance of eye movement measures between formats. The correlations between the dependent variables are all within the range of $r = -0.4$ and $r = 0.9$. Levene's test for equality of variances shows that there is homogeneity for all dependent variables

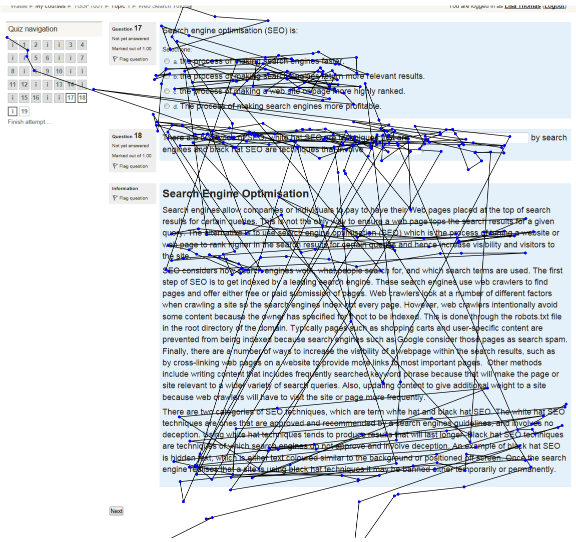

FIGURE 7. **Example of answer-seeking behaviour.**

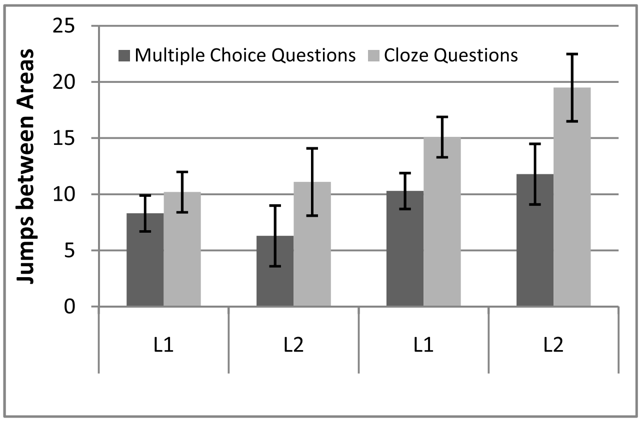

FIGURE 8. **Jumps between questions and text areas for each reader type, format and question type.**

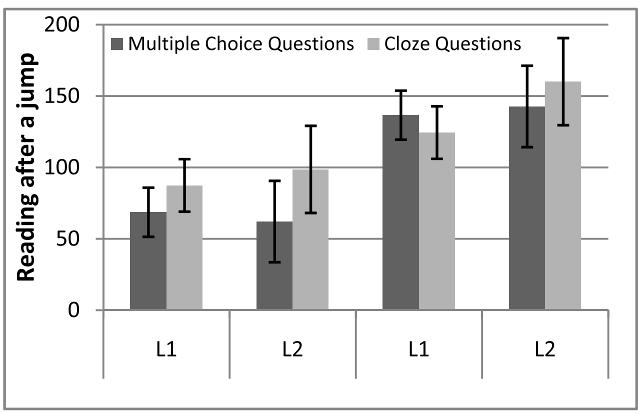

FIGURE 9. **Reading measured after a jump for each reader type, format and question type. (A: $T \rightarrow T/Q$; B: T/Q; C: $T \rightarrow Q$; D: $Q \rightarrow T \rightarrow Q$).**

(significance $>$ 0.05) and the Box's M value of 17.092 ($p = 0.207$) is interpreted as non-significant so we can be satisfied that there is homogeneity in the variance-variance-covariance matrices.

There is a significant difference in eye movement measures based on format, $F(4.23) = 4.199$, $p = 0.011$; Wilk's $\lambda = 0.578$, partial $\eta^2 = 0.422$, but no significant difference between the L1 and L2 reader and or effect of interaction.

Between-subjects ANOVAs are used to determine how the eye movements differ for the formats. Format has a significant effect the amount of reading after jumps from multiple-choice questions ($F(1, 26) = 10.006$, $p = 0.004$), and jumps between cloze questions and text ($F(1, 26) = 7.050$, $p = 0.013$), but not on jumps between multiple-choice questions and text or the amount of reading after jumps from cloze questions jumps.

This analysis partly confirms the hypothesis that participants would on average show less answer seeking behavior when presented with format A compared to B. Although pre-reading the text before seeing the questions does not decrease the amount of answer seeking observed for the multiple-choice questions it does for the cloze questions, which confirms the hypothesis for that kind of question.

There is no significant difference between the L1 and L2 readers in answer-seeking behavior. This contrasts with the other analyses in this paper where a clear difference was evident. This is an important finding: as the aim is to level the playing field for readers, and providing questions and text together will do so.

V. DISCUSSION & RECOMMENDATIONS

The overarching question of this study is whether the presentation of text and questions affects performance outcomes and how students' implicitly interact with the learning materials. We also investigate the effects of presentation on two groups of readers, L1 and L2 readers. The availability of learning materials to a wide and varied audience is becoming more common with the growth of online learning environments and online courses, such as MOOCs. More learners are reading materials written in their non-native language. The effects of this need to be explored further, the importance is only growing as accessibility of English MOOCs available to non-English (L2) readers.

The results from this study generally confirm that whilst L2 readers take longer to read context their comprehension is no different to L1 readers [16]. Delving deeper into why this discrepancy in read times occur we move to eye movements for insight. When reading text with no questions present L2 readers are observed to have higher numbers of fixations for longer duration than L1 readers. The divide between L1 and L2 readers decreases to being primarily due to fixation duration once the questions and text are presented together.

Fixation duration has been linked to comprehension and is an indicator for processing of words. Uncommon words often require longer fixations as do out of place words. It is plausible that L2 readers have a less comprehensive English vocabulary than L1 readers. The lesson to be learnt is that choice of vocabulary in writing learning materials is important whereby use of colloquial and uncommon words or phrases should be avoided or minimized so that

L2 readers are not impacted. Furthermore, time taken to complete a task should only be used as a performance measure if L1 and L2 readers are divided.

The study confirmed the hypothesis that the presentation format affected the L1 and L2 participants in the same way. Additionally, there is no significant difference between the L1 and L2 readers in answer-seeking behavior. This is an important finding as it means that any conclusions drawn regarding how presentation format affects students can be generalized for both reader types and it is not an additional factor that creators of learning materials have to take into account.

Indeed the hypothesis that presentation format will affect outcomes and eye movements were confirmed. The formats elicited distinct eye movement and reading behaviors. The format can therefore be manipulated to promote specific behaviors. In the next subsection recommendations are made based on these observations. There were some surprising results on the effects of presentation format that will now be discussed.

Foremost, the performance results from format D ($Q \rightarrow T \rightarrow Q$) are somewhat surprising given that pretesting with multiple-choice questions has been shown to benefit subsequent learning [5]. Memorizing pre-test questions instead of answering them has shown to improve recall of correct answers of the pre-exposed questions [5]. This is analogous to format D where participants were given the comprehension questions before they read the material to answer them. Participants were told that they should read the questions and were welcome to answer them if they wished. Yet the study showed no improvement in comprehension score compared to the control presentation format which required participants to rely purely on their memory of the text to answer the questions (format C: $T \rightarrow Q$).

Furthermore, format A ($T \rightarrow T/Q$) had two surprising effects on participants' behavior; first was that for this format there was no correlation between participants' quiz scores and their subjective ratings of understanding. This is surprising because for formats C and D ($Q \rightarrow T \rightarrow Q$) where the participants had to answer the questions without the text being available there were very strong correlations. The effect of showing the text before asking the questions was believed to at least partially mimic these formats thereby partially enhancing the ability to subjectively rate ones understanding. This however was not the case. Secondly, it was hypothesized that participants would read format C more thoroughly, in terms of fixations and reading ratio, than format A. This was not found either; instead participants read the text in format A as thoroughly as participants did in format C.

A. RECOMMENDATIONS FOR PRESENTING TEXT AND ASSESSMENT QUESTIONS

The observations from this study imply recommendations for educators designing courseware in an eLearning environment as well as design considerations for developers of eLearning environments. We have established that the presentation of text and questions impacts learning outcomes and reading behavior. The presentation format can be manipulated to optimize the performance outcomes of students, thereby increasing their understanding.

Formats C and D promote accurate self-assessment of comprehension, which in turn minimizes both under- and over-estimation of knowledge. Formats A and C promote thorough reading of the learning materials compared to D, therefore in the context of learning this is a more optimal outcome. Given the aims promote thorough reading and accurate self-assessment format C is optimal.

The differences in eye movement measures and reading behaviors reflect the overall purpose and goals placed on the reader. If an educator wants to promote thorough reading, the goals should not be targeted with the use of questions. In this case, students only read the parts of the text that they think contains the answers. However, not showing the text with the questions means that the students have to rely too heavily on short term memory which impacts their scores. The happy medium is format A where the students are requested to read the text and then move on to answer the compression questions. This raises the question of how to make students read the text before moving on to the questions and text page. This is where eye tracking can be utilized. The eye tracker can be integrated into the learning environment so that it can monitor reading behavior. Once the student has read the text then the learning environment would allow the student to move on to the questions.

These observations can also be used in the design of adaptive eLearning systems. Whilst we have established that the formats affects L1 and L2 readers in the same way, so the same learning materials can be used for both reader groups, this does not mean the format has to stay the same throughout the learning process. With the integration of eye tracking into the eLearning environment the reading behavior can be used to drive adaption to a student's current state. One format may not always be appropriate for them. If the student is tired and unmotivated, trying to get them to read the material thoroughly may not be the best idea. Instead they can be given format B (T/Q) so that they glean the surface knowledge and the deeper materials can be stored for when they log on next. Alternatively, since students often start motivated, format C can at first be shown to them, and as they start to lose that motivation formats A and then B can be shown to them.

Additionally, answer-seeking behavior can be used to detect how easy or difficult a student find the text. In previous studies we found that there are ranges of answer-seeking behavior for each question as well as between students [55], [56]. 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. If a student is found to be having no 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. Furthermore, this information can be reported back to the author of the learning materials so they can gauge the difficulty of the questions and improve them based on this feedback.

VI. CONCLUSION AND FURTHER WORK

This study was designed to increase our understanding of how text and questions presented in eLearning environments can affect eye movements and performance outcomes. These effects are investigated for L1 and L2 readers. We found that presentation of text and questions affect L1 and L2 readers consistently. There is a difference between L1 and L2 readers in that L2 readers take longer to complete the task however they are otherwise no different. Following this observation we observe that L2 readers have consistently longer fixation durations and in the situation where reading is the primary task, also L2 reader have more fixations than L1 readers.

Making participants rely on memory to answer assessment questions promotes more accurate subjective ratings of understanding. When participants are asked assessment questions after reading the content and have no reference back to the text they can more accurately gauge their understanding. When shown the text with the assessment questions participants are unable to gauge their own understanding.

The observations from the study show that different presentation sequences of text and questions affect performance outcomes and eye movements of participants. The order in which text and assessment questions are presented to students can therefore be manipulated to optimize performance outcomes and reading behavior.

A limitation of this study is that only two types of questions that are investigated in this analysis, being multiple choice and cloze questions. These are commonly used question type but not the only types available in eLearning environments, so further research should include to see what effect other question types have on the observed behavior.

Further exploration on how presentation formats affect behavior on mobile devices would be beneficial given the prevalence of this technology, as this study only considers reading from a computer screen in a university setting.

An area of interest is to investigate the relationship between eye movements and subjective comprehension. Although we touch on it in this investigation we want to investigate this relationship further. In a follow-up study we propose recording eye gaze from participants as they once again read through a tutorial and quiz. After each tutorial slide we can ask participants for their subjective comprehension score as well as other factors regarding how they read and perceived the text.

ACKNOWLEDGMENT

The authors would like to thank the participants of the study. Thank you to the reviewers for your valuable feedback and to all of the people who proofread drafts.

REFERENCES

[1] D. Romer, "Do students go to class? Should they?" *J. Econ. Perspect.*, vol. 7, no. 3, pp. 167–174, 1993.

[2] R. Woodfield, D. Jessop, and L. McMillan, "Gender differences in undergraduate attendance rates," *Stud. Higher Edu.*, vol. 31, no. 1, pp. 1–22, 2006.

[3] H. Kang, "Understanding online reading through the eyes of first and second language readers: An exploratory study," *Comput. Edu.*, vol. 73, pp. 1–8, Apr. 2014.

[4] R. C. Clark and R. E. Mayer, *e-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning*. New York, NY, USA: Wiley, 2011.

[5] J. L. Little and E. L. Bjork, "Pretesting with multiple-choice questions facilitates learning," in *Proc. Annu. Meeting Cognit. Sci. Soc.*, 2012, pp. 1–6.

[6] K. Rayner, "Eye movements in reading and information processing: 20 years of research," *Psychol. Bull.*, vol. 124, no. 3, pp. 372–422, Nov. 1998.

[7] L. Frazier and K. Rayner, "Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences," *Cognit. Psychol.*, vol. 14, no. 2, pp. 178–210, Apr. 1982.

[8] K. Rayner, K. H. Chace, T. J. Slattery, and J. Ashby, "Eye movements as reflections of comprehension processes in reading," *Sci. Stud. Reading*, vol. 10, no. 3, pp. 241–255, 2006.

[9] R. E. Bohn and J. E. Short, "How much information? 2009 report on American consumers," Global Inf. Ind. Center, Univ. California, San Diego, San Diego, CA, USA, Tech. Rep., 2009.

[10] A. Dillon, "Reading from paper versus screens: A critical review of the empirical literature," *Ergonomics*, vol. 35, no. 10, pp. 1297–1326, 1992.

[11] A. Dillon, *Designing Usable Electronic Text: Ergonomic Aspects of Human Information Usage*. Boca Raton, FL, USA: CRC Press, 2004.

[12] B. Sparrow, J. Liu, and D. M. Wegner, "Google effects on memory: Cognitive consequences of having information at our fingertips," *Science*, vol. 333, no. 6043, pp. 776–778, Jul. 2011.

[13] D. DeStefano and J.-A. LeFevre, "Cognitive load in hypertext reading: A review," *Comput. Human Behavior*, vol. 23, no. 3, pp. 1616–1641, May 2007.

[14] A. Dillon and R. Gabbard, "Hypermedia as an educational technology: A review of the quantitative research literature on learner comprehension, control, and style," *Rev. Edu. Res.*, vol. 68, no. 3, pp. 322–349, 1998.

[15] A. Mangen, B. R. Walgermo, and K. Brønnick, "Reading linear texts on paper versus computer screen: Effects on reading comprehension," *Int. J. Edu. Res.*, vol. 58, pp. 61–68, Jan. 2013.

[16] A. J. Rockinson-Szapkiw, J. Courduff, K. Carter, and D. Bennett, "Electronic versus traditional print textbooks: A comparison study on the influence of university students' learning," *Comput. Edu.*, vol. 63, pp. 259–266, Apr. 2013.

[17] Y. J. Rho and T. D. Gedeon, "Academic articles on the Web: Reading patterns and formats," *Int. J. Human-Comput. Interact.*, vol. 12, no. 2, pp. 219–240, 2000.

[18] K. O'Hara and A. Sellen, "A comparison of reading paper and on-line documents," in *Proc. ACM SIGCHI Conf. Human Factors Comput. Syst.*, Atlanta, GA, USA, 1997, pp. 335–342.

[19] L. Anderson-Inman and M. A. Horney, "Supported eText: Assistive technology through text transformations," *Reading Res. Quart.*, vol. 42, no. 1, pp. 153–160, Jan./Feb. 2007.

[20] D. McKay, "A jump to the left (and then a step to the right): Reading practices within academic ebooks," in *Proc. 23rd Austral. Comput.-Human Interact. Conf.*, Canberra, Australia, 2011, pp. 202–210.

[21] C. C. Marshall and S. Bly, "Turning the page on navigation," in *Proc. 5th ACM/IEEE-CS Joint Conf. Digit. Libraries (JCDL)*, Jun. 2005, pp. 225–234.

[22] D. Beymer and D. M. Russell, "WebGazeAnalyzer: A system for capturing and analyzing Web reading behavior using eye gaze," in *Proc. CHI Extended Abstracts Human Factors Comput. Syst.*, 2005, pp. 1913–1916.

[23] M. Bernard and M. Mills, "So, what size and type of font should i use on my website?" *Usability News*, vol. 2, no. 2, pp. 1–5, 2000.

[24] D. Beymer, D. Russell, and P. Orton, "An eye tracking study of how font size and type influence online reading," in *Proc. 22nd Brit. HCI Group Annu. Conf. People Comput., Culture, Creativity, Interact.*, vol. 2. Liverpool, U.K., 2008, pp. 15–18.

[25] J. S. Mansfield, G. E. Legge, and M. C. Bane, "Psychophysics of reading. XV: Font effects in normal and low vision," *Invest. Ophthalmol. Vis. Sci.*, vol. 37, no. 8, pp. 1492–1501, Jul. 1996.

[26] T. Traphagan, J. V. Kucsera, and K. Kishi, "Impact of class lecture webcasting on attendance and learning," *Edu. Technol. Res. Develop.*, vol. 58, no. 1, pp. 19–37, Feb. 2010.

[27] L. B. Breslow, D. E. Pritchard, J. DeBoer, G. S. Stump, A. D. Ho, and D. T. Seaton, "Studying learning in the worldwide classroom: Research into edX's first MOOC," *Res. Pract. Assessment*, vol. 8, no. 1, pp. 13–25, 2013.

[28] S. P. Liversedge and J. M. Findlay, "Saccadic eye movements and cognition," *Trends Cognit. Sci.*, vol. 4, no. 1, pp. 6–14, Jan. 2000.

[29] G. Underwood, A. Hubbard, and H. Wilkinson, "Eye fixations predict reading comprehension: The relationships between reading skill, reading speed, and visual inspection," *Lang. Speech*, vol. 33, no. 1, pp. 69–81, 1990.

[30] B. G. Prusty and C. Russell, "Engaging students in learning threshold concepts in engineering mechanics: Adaptive eLearning tutorials," in *Proc. Int. Conf. Eng. Edu. (ICEE)*, Belfast, Northern Ireland, U.K., 2011, pp. 1–10.

[31] C.-M. Chen, "Intelligent Web-based learning system with personalized learning path guidance," *Comput. Edu.*, vol. 51, no. 2, pp. 787–814, 2008.

[32] H. D. Surjono, "The design of adaptive e-learning system based on student's learning styles," *Int. J. Comput. Sci. Inf. Technol.*, vol. 2, no. 5, pp. 2350–2353, 2011.

[33] H. D. Surjono, "The evaluation of a Moodle based adaptive e-learning system," *Int. J. Inf. Edu. Technol.*, vol. 4, no. 1, pp. 89–92, 2014.

[34] T. J. Mehigan, M. Barry, A. Kehoe, and I. Pitt, "Using eye tracking technology to identify visual and verbal learners," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2011, pp. 1–6.

[35] D. Spada, M. Sánchez-Montañés, P. Paredes, and R. M. Carro, "Towards inferring sequential-global dimension of learning styles from mouse movement patterns," in *Adaptive Hypermedia and Adaptive Web-Based Systems*. Berlin, Germany: Springer-Verlag, 2008.

[36] Z. Liu, "Reading behavior in the digital environment: Changes in reading behavior over the past ten years," *J. Documentation*, vol. 61, no. 6, pp. 700–712, 2005.

[37] T. J. Mehigan and I. Pitt, "Intelligent mobile learning systems for learners with style," in *Tools for Mobile Multimedia Programming and Development*, D. Tjondronegoro, Ed. Hershey, PA, USA: IGI Global, May 2013, pp. 131–149.

[38] P. De Bra, "GRAPPLE: Learning management systems meet adaptive learning environments," in *Intelligent and Adaptive Educational-Learning Systems*. Berlin, Germany: Springer-Verlag, 2013, pp. 133–160.

[39] A. Y. Alsobhi, N. Khan, and H. Rahanu, "DAEL framework: A new adaptive e-learning framework for students with dyslexia," *Procedia Comput. Sci.*, vol. 51, pp. 1947–1956, Jun. 2015.

[40] S. T. Iqbal and B. P. Bailey, "Using eye gaze patterns to identify user tasks," in *Proc. Grace Hopper Celebration Women Comput.*, Chicago, IL, USA, 2004.

[41] C. S. Campbell and P. P. Maglio, "A robust algorithm for reading detection," in *Proc. Workshop Perceptive User Interfaces*, 2001, pp. 1–7.

[42] G. Buscher, A. Dengel, and L. van Elst, "Eye movements as implicit relevance feedback," in *Proc. Extended Abstracts Human Factors Comput. Syst. (CHI)*, Florence, Italy, 2008, pp. 2991–2996.

[43] S. Martinez-Conde, "Fixational eye movements in normal and pathological vision," *Prog. Brain Res.*, vol. 154, pp. 151–176, Sep. 2006.

[44] N. Nugrahaningsih, M. Porta, and S. Ricotti, "Gaze behavior analysis in multiple-answer tests: An eye tracking investigation," in *Proc. Int. Conf. Inf. Technol. Higher Edu. Training (ITHET)*, Oct. 2013, pp. 1–6.

[45] M.-J. Tsai, H.-T. Hou, M.-L. Lai, W.-Y. Liu, and F.-Y. Yang, "Visual attention for solving multiple-choice science problem: An eye-tracking analysis," *Comput. Edu.*, vol. 58, no. 1, pp. 375–385, 2012.

[46] S.-C. Chen, H.-C. She, M.-H. Chuang, J.-Y. Wu, J.-L. Tsai, and T.-P. Jung, "Eye movements predict students' computer-based assessment performance of physics concepts in different presentation modalities," *Comput. Edu.*, vol. 74, pp. 61–72, May 2014.

[47] C. Gütl *et al.*, "AdeLE (adaptive e-learning with eye-tracking): Theoretical background, system architecture and application scenarios," *Eur. J. Open, Distance, e-Learn.*, vol. 2, 2005. [Online]. Available: http://www.eurodl.org/materials/contrib/2005/Christian_Gutl.htm

[48] C. Calvi, M. Porta, and D. Sacchi, "*e5Learning*, an e-learning environment based on eye tracking," in *Proc. 8th IEEE Int. Conf. Adv. Learn. Technol. (ICALT)*, Jul. 2008, pp. 376–380.

[49] S. D'Mello, A. Olney, C. Williams, and P. Hays, "Gaze tutor: A gaze-reactive intelligent tutoring system," *Int. J. Human-Comput. Stud.*, vol. 70, no. 5, pp. 377–398, 2012.

[50] A. Paramythis and S. Loidl-Reisinger, "Adaptive learning environments and e-learning standards," in *Proc. 2nd Eur. Conf. e-Learn.*, 2003, pp. 369–379.

[51] A. Hyrskykari, P. Majaranta, A. Aaltonen, and K.-J. Räihä, "Design issues of iDICT: A gaze-assisted translation aid," in *Proc. Symp. Eye Tracking Res. Appl.*, 2000, pp. 9–14.

[52] J. L. Sibert, M. Gokturk, and R. A. Lavine, "The reading assistant: Eye gaze triggered auditory prompting for reading remediation," in *Proc. 13th Annu. ACM Symp. User Interface Softw. Technol.*, San Diego, CA, USA, 2000, pp. 101–107.

[53] P. Lach, "Intelligent tutoring systems measuring student's effort during assessment," in *Advances in Artificial Intelligence*. Berlin, Germany: Springer-Verlag, 2013, pp. 346–351.

[54] J. M. Fletcher, "Measuring reading comprehension," *Sci. Stud. Reading*, vol. 10, no. 3, pp. 323–330, 2006.

[55] L. Copeland and T. Gedeon, "The effect of subject familiarity on comprehension and eye movements during reading," in *Proc. 25th Austral. Comput.-Human Interact. Conf., Augmentation, Appl., Innov., Collaboration*, 2013, pp. 285–288.

[56] L. Copeland and T. Gedeon, "Measuring reading comprehension using eye movements," in *Proc. IEEE 4th Int. Conf. Cognit. Infocommun. (CogInfoCom)*, Dec. 2013, pp. 791–796.

[57] G. W. McConkie and K. Rayner, "The span of the effective stimulus during a fixation in reading," *Perception Psychophys.*, vol. 17, no. 6, pp. 578–586, 1975.

[58] J. Hyönä, R. F. Lorch, Jr., and M. Rinck, "Eye movement measures to study global text processing," in *The Mind's Eye*, J. Hyönä, R. Radach, and H. Deubel, Eds. Amsterdam, The Netherlands: North Holland, 2003, ch. 16, pp. 313–334.

[59] J. Dunlosky and A. R. Lipko, "Metacomprehension: A brief history and how to improve its accuracy," *Current Directions Psychol. Sci.*, vol. 16, no. 4, pp. 228–232, 2007.

LEANA D. COPELAND received the B.Sc. and B.I.T. (Hons.) degrees from Australian National University, Canberra, Australia, in 2012, where she is currently pursuing the Ph.D. degree in computer science. Her current research interests are human–computer interaction, eye tracking, online education, and bioinspired computing.

TOM D. GEDEON received the B.Sc. and Ph.D. degrees from the University of Western Australia. He was the Deputy Dean and Head of Computer Science with Australian National University (ANU). He is currently the Chair Professor of Computer Science with ANU, and was the President of the Computing Research and Education Association of Australasia. He was the President of the Asia-Pacific Neural Network Assembly, has been the General Chair of the International Conference on Neural Information Processing twice, has been nominated for VC's awards for post-graduate supervision at three universities, and serves on a number of journal advisory boards as a member, an Associate Editor, or an Editor.

His research focuses on bioinspired computing (mainly neural, fuzzy, and evolutionary) and human-centered computing (mainly eye gaze and physiological signals) to construct humanly useful information resources (hierarchical and time series knowledge), and social good (medical and educational) applications.