The Reading Matrix: An International Online Journal
Volume 20, Number 1, April 2020

Online Research Strategies of L2 Readers: Evaluating Strategic Competence through Mixed Methods

Caleb Prichard
Okayama University

Andrew Atkins
Kindai University

ABSTRACT

Researching from multiple-documents online is an essential activity, but it has generally been overlooked by second language (L2) educators. This study uses navigation tracking, eye tracking, a post-reading recall task, and interviews to examine the degree to which Japanese L2 readers maintain task awareness when researching online. Key reading variables analyzed were attention to relevant versus irrelevant pages and repeated versus novel content. Finally, the experiment analyzed the connection between the reading measures and post-reading recall scores. The findings suggest that the readers displayed moderate competence overall, but some participants lacked strategic competence. There was a significant correlation between reading relevant content and recall performance. The research has several implications for L2 instructors.

INTRODUCTION

Second language (L2) reading instruction, testing, and research all tend to focus on the comprehension of single texts (Karimi, 2015) despite the fact that reading multiple documents online is the norm for a wide range of information-gathering tasks, from planning a vacation to completing school assignments (Bråten & Strømsø, 2011). Researching from multiple texts involves different skills and strategies than reading single documents, such as: searching for, sifting through, and selecting appropriate texts; processing the selected passages by reading, skimming, or scanning; and, evaluating and synthesizing the information (Naumann, Richter, Christmann, & Groeben 2008).

Considering the mass of content on the internet and the ease with which hyperlinked texts can be accessed, metacognitive strategies for planning and maintaining focus on task-relevant materials are particularly necessary (Bråten & Strømsø, 2011; Karimi, 2015). While much of the L2 research on metacognitive skills has relied on surveys (Prichard & Atkins, 2019), several L1 studies have examined this issue by tracking access to relevant and irrelevant pages (e.g., OECD, 2011; Rodicio, 2015). Online measures (i.e., real-time measures, such as navigation tracking and eye tracking) are common in L1 reading research because such data can reliably verify learners' reading strategies on a moment-to-moment basis given a specific task. However, while L2

research using eye tracking is quickly gaining momentum (Godfroid, 2019), research utilizing online measures on multiple-document L2 reading is lacking.

Considering this gap, this experiment utilizes navigation tracking and a post-reading recall task to evaluate the English research strategies of Japanese L2 learners. Eye tracking and interviews are used for data triangulation. The study primarily focuses on the degree to which learners consider relevance and maintain task awareness when selecting and reading texts online. The study also examines the relationship between reading behavior and post-reading task performance. Ultimately, the research aims to examine whether multi-document reading online is an area worthy of deeper consideration among ESOL educators.

LITERATURE REVIEW

Navigation skills and media literacy are now essential for the digital age. Reading online provides benefits over print-based reading as internet searches instantly retrieve millions of texts. Moreover, hypertexts, non-linear individual pages linked to each other, can enhance complex learning compared to printed materials since hypertexts provide learners with flexibility in accessing and organizing information based on one's needs (Naumann, et al., 2008). Strategically processing multiple texts is closely related to task performance in digital reading since meaning is constructed by effectively selecting and navigating through texts (Salmerón, Cañas, Kintsch, & Fajardo, 2005).

However, the mass of information, unreliable or irrelevant content, and the non-linearity of hypertexts often overloads or distracts readers (Coiro, et al., 2008; OECD, 2011; Naumann, et al., 2008). Therefore, many readers have difficulty navigating the web to research needed information (Fajardo & Villalta, 2016; Kuiper, et al., 2005). Effectively selecting articles, navigating through and processing the text, and formulating conceptual and lexical connections between pages is extremely challenging (Naumann, et al., 2008). This study is concerned with two key areas that can minimize these challenges: first, and primarily, text selection of relevant sources, and second, effectively dealing with overlapping information.

Text Selection for Task-based Reading

Considering that search engines typically yield millions of hits, many irrelevant and unreliable, effective online researchers select texts based on their utility. Blindly selecting the top links from the search results page can often be effective, but it can lead to poor outcomes (Rodicio, 2015). An effective search depends on coming up with effective search terms and evaluating relevance and reliability of the results by checking the title, URL, and the article snippet provided. After clicking on a link to a selected article, readers may preview the text to continue to evaluate its overall utility (Bråten & Strømsø, 2011). Processing multiple texts effectively involves maintaining task awareness and navigation skills as users scroll and click links to process texts in needed ways, such as previewing, skimming, scanning, careful linear reading, and/or re-reading (Naumann, et al., 2008).

Because the challenges of effectively researching a topic from multiple documents, metacognitive strategies to plan and monitor one's reading are essential (Naumann, et al., 2008). Selective attention, which is the capacity to focus on relevant content (Stevens & Bavelier, 2012), is vital considering the huge number of search results, the number of unreliable sites, and

the seemingly infinite number of access points and links (Smallwood et al., 2007). Studies have shown that more effective readers "learned and recalled more important information… because they were more metacognitively aware of how and when" to use selective attention (Reynolds, Shepard, Lapan, Kreek, & Goetz, 1990, p. 749). Selective attention is not only a metacognitive strategy; it also reflects one's cognitive ability to ignore irrelevant content (Tipper & Baylis, 1987), especially s*eductive details* that tend to distract readers (Peshkam et al., 2011). Because processing written information uses up working memory (a finite capacity to hold information in one's mind for behavior and decision-making), maintaining focus on task-relevant content increases cognitive efficiency (McCrudden et al., 2011).

L1 Studies on Task Relevance. Think-aloud research, which involves participants vocalizing their thinking while performing a task, has suggested that readers who can better recognize relevant texts are more successful in essay writing (Anmarkrud, McCrudden, Bråten, & Strømsø, 2013). Beyond self-report protocols, L1 researchers have observed and analyzed participants' attention to relevant content by tracking page navigation. The studies tend to show that readers more often access and read closely task-relevant pages, but this is not always the case (Cress & Knabel, 2003; OECD, 2011; Rodicio, 2015). While readers often clicked relevant pages, they frequently relied on page rank and sometimes clicked on links based on misleading keywords (Rodicio, 2015).

The research also examined the effect of the strategies used. Overall, the findings highlight the importance of task-oriented navigation and metacognitive reading strategies in that better performing participants more often read crucial pages and avoided viewing irrelevant texts (OECD, 2011; Rodicio, 2015). In each of the 19 countries examined by OECD (2011), students (aged 15) differed in their digital reading performance based on the number of task-relevant pages accessed. This was true even when accounting for their print reading proficiency; in fact, the number of task-based pages viewed was more of a factor than print reading ability.

Strategic Processing of Selected Texts

In researching from multi-documents, readers must not only seek relevant and reliable information, they also need to effectively comprehend, consider, store, and recall it (Carenini, Cheung, & Pauls, 2013). If there is conflicting or inconsistent information, readers must recognize this (White et al., 2002) and use cross-text elaboration or accumulation strategies depending on their goal (Bråten & Strømsø, 2011). Cross-text elaboration is to compare and contrast, synthesize, and evaluate overlapping information, while accumulation is gathering as much information as possible. Although not discussed in the work of Bråten and Strømsø (2011), skipping content already adequately comprehended and processed from another source could be a useful strategy. This is because reading such previously read content could waste reading time and working memory that could better be utilized to identify, process, and consider novel content.

Overall, research using think-aloud protocols has shown that effective readers of multiple documents are very active, making connections between texts and often return to re-read relevant texts (Bråten & Strømsø, 2011). On the other hand, less effective readers tend to read more linearly, making little mental connections between texts.

L2 Research

Numerous studies have evaluated the use and effect of metacognitive strategies in L2 reading using survey protocols, such as The Metacognitive Online Reading Strategies Questionnaire (MORS-Q; Romly, Badusah, & Maarof, 2017). The planning construct which concerns previewing and selective attention, has predicted L2 reading proficiency in Thailand (Phakiti, 2003), China (Zhang & Seepho, 2013), Iran (Ghafournia & Afghari, 2013), and Malaysia (Romly, Badusah, & Maarof, 2017). Other researchers have used the SORS (Sheorey & Mokhtari, 2001), and its global construct, which considers skills such as previewing and considering one's purpose, has been shown to correlate with L2 reading proficiency (Zhang & Wu, 2009; Prichard, 2014). However, such survey results concern participants' general reading tendencies, not strategies used on actual multi-document research tasks.

Karimi (2015, 2017) is one researcher who has specifically examined multiple-document reading. His 2015 study looked at the relationships between L1 reading ability, L2 reading ability on single and multiple texts, and reported strategy usage among Iranian learners of English. The results suggested that strategy usage contributed more to L2 comprehension of multiple texts than L1 proficiency did. Moreover, strategic competency was more related to success in multiple-text L2 reading than in single-test L2 reading.

While the above L2 research is insightful and strongly hints at the importance of strategic competence, the studies have relied on self-report protocols, which have numerous limitations. Participants often misreport their strategy use (Bax & Weir, 2012; Prichard & Atkins, 2016) since they may try to give the ideal response or they may be unconscious of their strategies (Denscombe, 2014). Using online measures, such as navigation tracking or eye tracking, can verify and greatly deepen the findings as researchers can reliably observe actual behaviors and decisions made by readers. While eye tracking is getting much more common in L2 research, such methods are still relatively lacking in research on the global strategies of learners.

However, Prichard and Atkins (2019) did examine Japanese L2 English readers' global strategies using eye tracking, albeit of single documents. The results showed that participants tended to focus more on relevant content but almost all read the whole page, even the task-irrelevant sections. This supported observations by Hirata and Hirata (2010) that Japanese university-aged learners often struggle accessing English media strategically, lacking ability to preview or effectively navigate English pages.

Although Japanese demonstrated strategic competence in task-based, multiple-document L1 reading in the OECD data (2011), why does the research insofar suggest they often lack strategic competence to do this in the L2? One reason may be explained by research showing that L2 ability can mediate L1 strategic competency in the L2 (Guo, 2018). Since the English reading proficiency of Japanese is the lowest amongst its peers (including east Asian counterparts, Korea and China, and G8 Industrialized Nations; ETS, 2017), limited English reading proficiency may hinder Japanese ability to strategically research online.

It has also been argued that Japanese may lack global reading strategies in English partially because of the pedagogy of English education in Japanese junior high and high schools (Prichard & Atkins, 2019). Grammar-translation is still a common method in Japan, and this may hinder metacognitive strategy development in the L2 (Sakurai, 2015). In grammar-translation, texts are read word for word from the beginning to end, and selective attention is not given to any particular section. In addition, reading texts are typically assigned, so students do not need to consider text relevance. Japanese learners of English tend to lack opportunities to read multiple

documents to research or for task-based reading, and such learners may not acquire strategic reading habits in the L2 (Zhang, 2010).

METHODS

This study aimed to evaluate the online reading strategies of Japanese L2 learners of English when researching a specific topic from a search engine results page. After considering the research task, participants were provided a replica search results page linking to eight pages, four relevant and four irrelevant to the task. The study revolved around the two primary research questions:

1. Do advanced Japanese L2 readers of English pay selective attention to task-relevant articles in internet search engine results?
2. Does attention to task-relevant pages predict ability to recall task-relevant content?

A secondary purpose was to evaluate how readers dealt with content repeated on multiple pages. Do participants read or avoid such repeated content? Does this influence task-performance? Answering these questions has significant implications for L2 reading educators.

This study replicated aspects of Rodicio (2015), an L1 study, in that half the websites were relevant and irrelevant to the research task. Reliability of the content was removed as an additional variable, as it was judged beyond the scope of this paper. This study was original as it involved L2 learners. Moreover, eye tracking and interviews were used to verify and further explore the data, as explained below. Other differences from the Rodicio study are described below.

Participants

The participants included 24 second-year students from various faculties of a Japanese university. They were enrolled in a required academic reading and writing course with the main objective of developing the ability to write an effective research paper. All participants had scores of at least 650 on the Test of English for International Communications (TOEIC). Their proficiency may be considered advanced for Japanese university students.

Equipment

A Gazepoint GP3 HD Eye Tracker with a 150Hz refresh rate was utilized to triangulate the results. The equipment recorded screencasts of the participants' navigation behaviors along with their eye movements. Because the research analyzed fixations within large sections of the screen (whole pages or sections rather than words or sentences), the accuracy and refresh rate of the tracker was deemed appropriate. The participants could read in an authentic setting as chin rests were not involved and the eye tracker was placed under the screen on a small tripod. Participants' eyes were roughly 60 centimeters from the 24-inch high-definition monitor.

Materials

Participants were shown mock results from a search engine page, as in Rodicio (2015). Here the results were for "playing year round sports" for their research task, which was to find

disadvantages of youth playing year round sports. The topic was selected after piloting among students at the institution suggested that participants were not familiar with the specific topic, but interested in it. In the piloting, students in a previous cohort of the class read about a series of controversial issues, and answered open ended discussion questions about the content. Year-round youth athletics was the only topic in which none of the students indicated they had previously read about.

The search page linked to eight articles, with four relevant and four irrelevant to the specific research task. The task-relevant articles were evenly distributed throughout the results page (Articles 2, 3, 5, 8), as shown in Table 1. Two of the irrelevant articles mentioned the opposite of the research task, and two were about the general topic, but not the specific topic.

Table 1. The Content of the Articles

Article #	Content Relevance
1	Irrelevant (advantages, not disadvantages)
2	Relevant: injury risk
3	Relevant: burnout, injury risk, social development hindrance
4	Irrelevant (only loosely related to the general topic)
5	Relevant: injury risk
6	Irrelevant (advantages, not disadvantages)
7	Irrelevant (only loosely related to the general topic)
8	Relevant: burnout, injury risk, social development hindrance, less skill development

The relevant articles varied in how much repeated and novel information they contained. Article 5 described only one main idea, which was already discussed in Articles 2 and 3. Articles 3 and 8 described several relevant points, but each contained only one novel point not mentioned in an article presented above it. The articles chosen were reviewed for relevance by the two authors and two native speaking graduate students. All agreed on the task-relevant sections.

The search results titles, hyperlinked in blue just as they appeared on Google (see Figure 1), clearly suggested whether or not the articles would discuss points related to the task. The titles in the search page for Article 2 and 5 revealed that the content was basically the same for these articles. The search page also showed the URL and a snippet of the text, taken from actual Google search results.

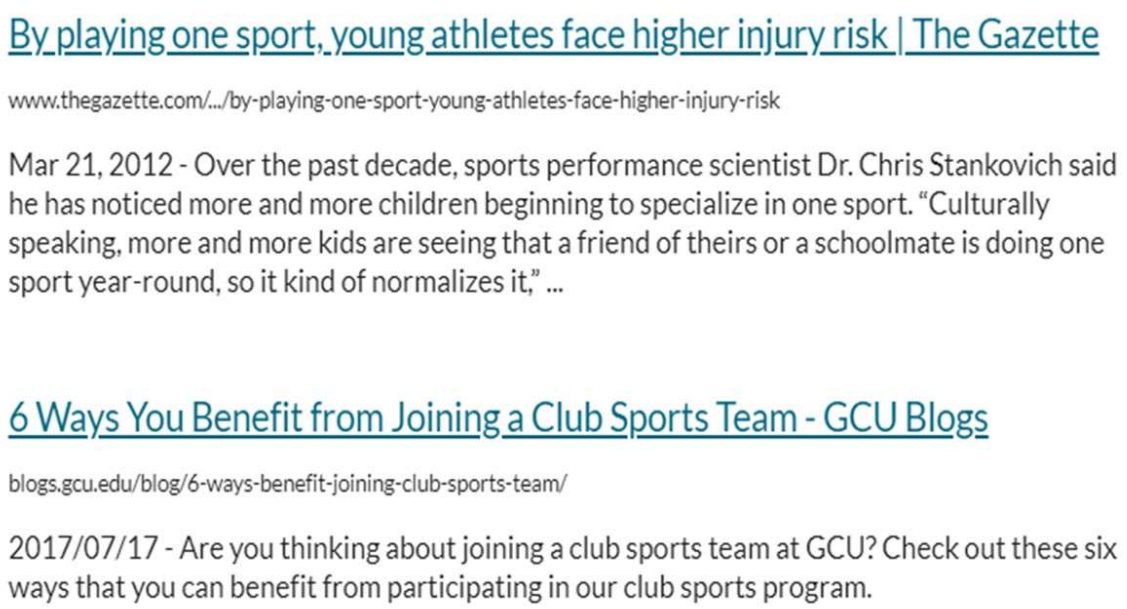

Figure 1. Screenshots of the search engine page showing Article 5 and Article 6.

The texts were edited from the original sources in order to empirically examine the variable of task-relevance. Each of the eight articles was revised so that it had exactly 650 words. In Rodicio (2015), the texts were shorter (mean word count was 445.20), but there were more articles (12).

Because vocabulary coverage is known to affect reading, 96-98% of the running words in each section were in the top 3,000 words (excluding loanwords and pronouns) in the British National Corpus (2007). Articles were edited so that their readability was similar, based on the readibility.io program (2018). The layout of the pages was the same, including the text size and font. Each page contained an equally-sized photo underneath the title.

In sum, because the content, language, and layout were carefully controlled, any disparities in the access times on the relevant and irrelevant pages could be mainly attributed to the metacognitive reading strategies of the participants, not to text difficulty or novelty.

Procedures

Students in one of the authors' courses were required to sign up for a time to visit the professor's office for an out-of-class assignment. The period was within the first two weeks of the course before any reading strategies had been introduced or practiced. During their appointed time, each participant first was asked to answer six pre-reading questions on the topic they were to research to ensure they were fully focused on the task before reading. The students then were asked to read the following directions about the specific research task:

> You will soon be shown a Google search page showing results for "playing sports year round." You have 15 minutes to research the disadvantages of playing year round sports. Click the links and the back button, but you cannot access other pages or a dictionary. Note taking is not allowed. Afterwards, you need to write on paper as many

> disadvantages as you can. For each disadvantage, you need to write just one sentence explaining the disadvantage.

The participants then moved to a computer. If the students consented to participate in the study, their eyes were calibrated. (Other students finished the same task as it was a class assignment, but their online behavior was not recorded.) If there were no further questions, the reading task began.

After the 15-minute research task, the participants completed the recall task on paper. While Rodicio (2015) used nine multiple-choice questions, an immediate recall task was chosen for this study as it is generally considered a more valid and reliable measure, although not without limitations (Chang, 2006). The form indicated that the participants should note down the main disadvantages of year-round sports. They also needed to write at least one sentence explaining or detailing each point. The participants were given 15 minutes for the post-reading task.

After the post-reading task, 13 participants consented to take part in an interview for 20 minutes. Voices were audio recorded while the video recorded the screencast. During the interviews, participants viewed the screencast of their navigation and eye fixation map. Participants were asked to explain their rationale for any of the following behaviors displayed on the recording: clicking an article, skipping over an article, reading the entire article, skipping over a section, re-reading a section, and terminating reading the article.

Analysis

Pages accessed and view time on each page. The number of task-relevant and irrelevant pages accessed were compared. View times were tracked and compared on the four relevant and four irrelevant articles to evaluate participants' level of selective attention to the relevant pages. The time on the searched engine page was also tracked as more time on the page suggested closer evaluation of the articles.

Navigation and eye fixation data. For valid analysis, unique eye tracking approaches are needed depending on the specific reading task and research objective (Latif, 2019). Since relevant and irrelevant articles took up the whole page, access time on each page was used instead of eye tracking. The analysis of eye tracking was only used to examine if participants read different portions of the accessed texts. Eye fixation maps were evaluated to determine if participants read the whole section or the whole page. As in Prichard and Atkins (2019), *reading* (linear reading) was differentiated from previewing, skipping, scanning, or skimming, and was determined by three or more fixations on three or more lines from the top to the bottom. These reading behaviors were compared for relevant and irrelevant pages (relevant pages: Articles 2, 3, 5, 8) and for novel and repeated content (repeated content: Articles 5 & multiple sections of Articles 3 & 8, if the applicable content above it was already read). The two researchers coded this data independently. Inter-rater reliability was over 95%, reaching 100% after discussion.

Post-reading task. The number of task-relevant main ideas recalled counted as one point (maximum = 4). As the participants were asked to provide one supporting detail or example for each task-relevant main idea, another point was awarded (one per main idea) if this was accurately presented. The maximum number of points was therefore eight. A recalled unit was scored regardless of spelling and grammatical accuracy or whether it was verbatim or

paraphrased. The post-reading task was scored by the two researchers with an initial 96% inter-rater reliability rate.

Interviews. As mentioned above, while viewing their screencast and eye fixation map, 13 participants were asked to state their reading strategies and why they used them. The interviews were analyzed qualitatively based on pre-set and emergent codes, and the data were used to triangulate the quantitative data.

Results

Reading Strategies

Pages accessed. Of the eight articles available, the participants accessed a mean of 4.96 (*SD* = 1.49). In terms of the order of the accessed the articles, there were no two participants who followed the same path. Fourteen of the 24 participants (58%) first viewed Article 1 from the top of the search results though it was irrelevant to the task. The other ten participants accessed a relevant article first, especially Article 2 ($n = 6$). Overall, the relevant articles were accessed significantly more ($M = 3.25$, $SD = .74$) than the irrelevant articles ($M = 1.71$, $SD = 1.23$), $t(23) = 5.31$, $p < .0001$, $d = 1.38$. Although most participants accessed the Article 1, each of the relevant articles was accessed by more people (see Figure 2).

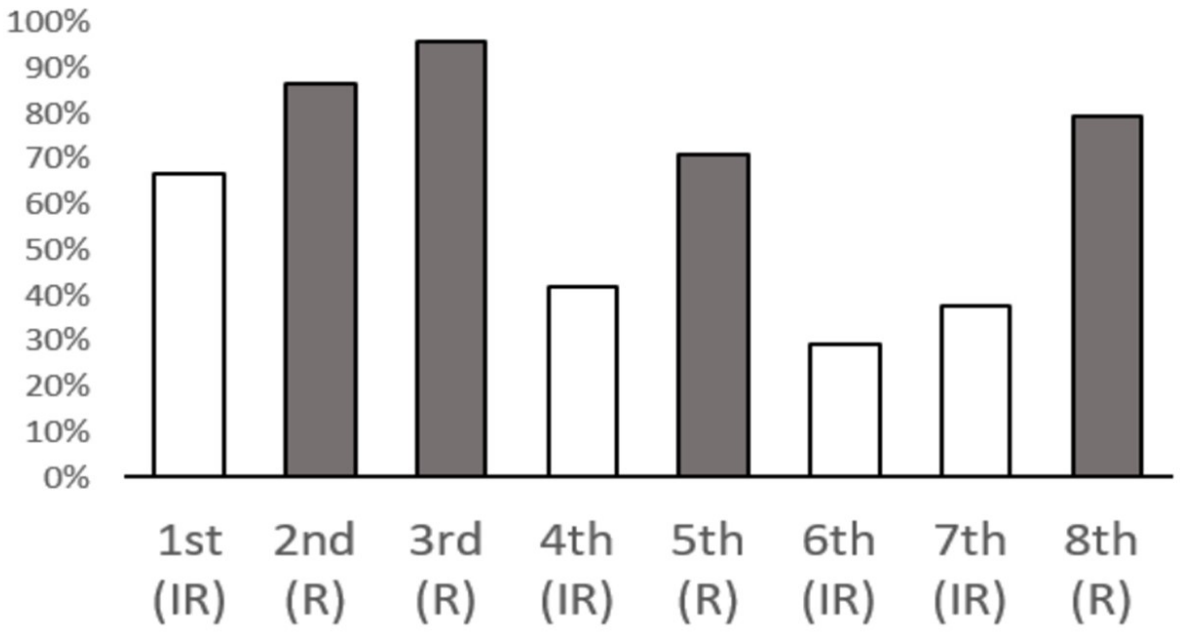

Figure 2. The percentage of participants who accessed each article. Note: IR = task-irrelevant article; R = task-relevant article.

Some articles were viewed multiple times by the same participant. Articles repeatedly viewed the most were the relevant articles that presented novel content, namely Article 3 (n = 13), Article 2 (n = 10), and Article 8 (n = 7).

Access time. The participants viewed the search result page for a mean of 25.92 seconds (SD = 31.20) before clicking on a link for the first time. Participants who first clicked to read the irrelevant article on the top viewed the search page for a significantly shorter time (M = 12.91, SD = 7.23) than those who skipped over the first article to click on a relevant article (M = 44.14, SD = 42.19), $t(22) = 2.74$, $p = .01$, $d = 1.13$. However, the total time participants spent on the search result page throughout the task did not have a significant relationship with the percent of relevant pages accessed, $r(22) = .22$, $p = .16$.

Considering the total task time, the participants spent much more time viewing relevant pages (71% of the total time) compared to irrelevant pages (17%) and the search results page (12%). They viewed the relevant articles significantly longer (M = 661.08 s, SD = 158.97) than the irrelevant articles (M = 160.38 s, SD = 165.01), $t(22) = 7.72$, $p < .0001$, $d = 3.09$. As is shown in Figure 3, one exception was Article 1, which was irrelevant but viewed a mean of 103.01 seconds (SD = 127.71).

Among the relevant articles, the two most viewed (3 & 8) each contained multiple task-relevant points. Article 5, which was relevant but did not present novel content, was viewed the least among the relevant articles, and even less than irrelevant Article 1.

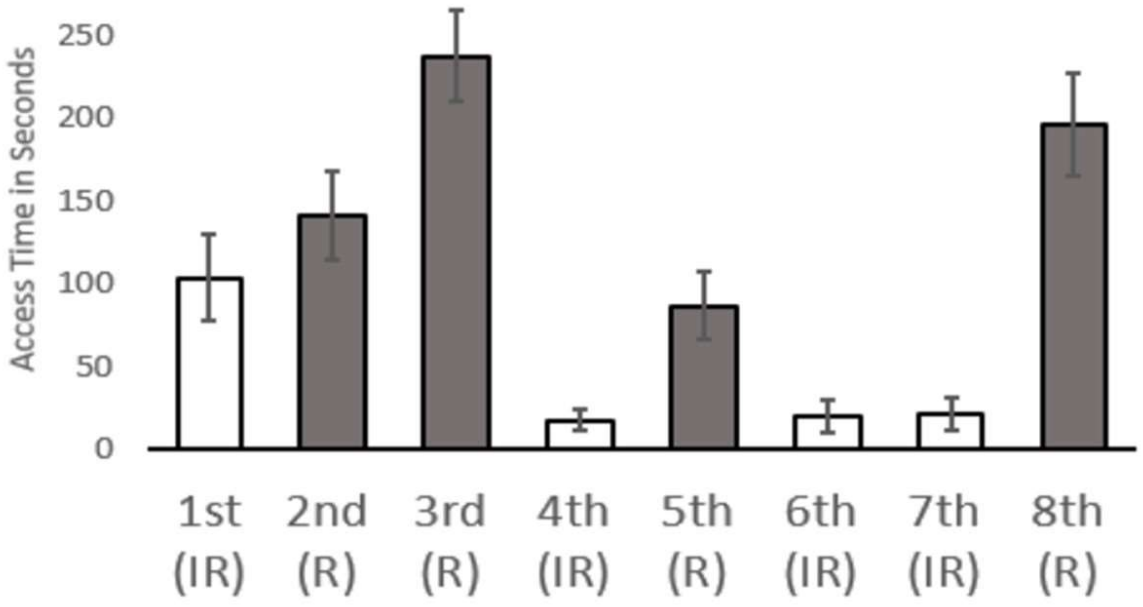

Figure 2. The mean access time in seconds on each article. Note: IR = task-irrelevant article; R = task-relevant article.

Eye fixation data. Eye movements confirmed that participants often left the article without reading the whole article, especially for the irrelevant articles. Although several participants (38%) accessed Article 7, which was only loosely related to the general topic, no one

read the whole article. However, among those who read the irrelevant Article 1, the whole article was read by roughly half of the participants (56.25%).

Among the relevant articles which had repeated content already presented in earlier articles, eye movements indicated most participants (article 3: 75%; article 8: 72%) did linearly read the repeated content as opposed to skipping over or scanning it.

Interview data. For the relevant articles, the 13 interview participants nearly always replied that they clicked on the article because it was related to what they needed to research. However, two participants reported clicking relevant Article 2 simply because it was next in order on the search page. They had read Article 1 before, and they had planned to read all articles in order. When inquired why they had read a relevant article which repeated content already read about in another article, the most common answer was they want to get more detailed information about the point. However, one replied he tends to always read everything.

When inquired about why they had read a relevant article a second or additional time, five answered that they wanted to clarify the content or that they wanted to review it so that they could better recall the information on the post-reading task. However, two indicated that they had clicked it by mistake since they forgot which articles they had read. One participant replied he had no reason for accessing an article a second time.

When asked why a relevant article was not read, two participants indicated that they had misread the link description. One said he had no reason for skipping the article. For relevant articles which had repeated content already read in another article, five participants indicated that they did not read it again because it was repeated content.

The participants gave a variety of answers for why they clicked on irrelevant articles. Eight answered that they thought they might be able to find relevant information even though they knew the article was mostly irrelevant. Six replied they had read Article 1 because it was at the top of the page. A few replied that they normally tend to read everything in the order presented. Four replied that they simply were not thinking or they clicked for no reason.

Post-task Performance and Its Relationship with Reading Strategies

The mean score on the post-reading recall task of the main points was 6.13 (SD = 1.45) out of eight. The time users spent on the search result page did not have a significant relationship with the total score, $r(22) = .18$, $p = .20$. The scores were also not significantly correlated with the ratio of relevant and irrelevant articles accessed, $r(22) = .05$, $p = .41$. Certain articles, especially relevant ones, were accessed multiple times. Even considering this, the ratio of clicking on relevant articles had no significant correlation with the recall score, $r(22) = .23$, $p = .14$. However, the total number of times each participant accessed the relevant articles (including to re-read or review) did have a significant moderate correlation with the recall score, $r(22) = .54$, $p < .0001$. In contrast, the total number of times participants accessed irrelevant articles did not have an effect, $r(22) = -.01$, $p = .49$.

The recall score significantly correlated with the ratio of access time on relevant versus irrelevant articles, $r(22) = .56$, $p < .0001$. The recall task score correlated with the total access time of relevant articles, $r(22) = .45$, $p = .01$, and it an even stronger reverse correlation on the irrelevant articles, $r(22) = -.60$, $p < .0001$.

While access time to relevant articles predicted task performance, time spent on relevant article 5, which repeated content from article 3, had a negative correlation with task recall score, $r(22) = -.34$, $p = .05$. The percentage of total reading time on articles that presented both relevant

and novel information (Articles 1, 3, 8) had a strong significant correlation with the recall score, $r(22) = .73$, $p < .0001$.

DISCUSSION

Did the L2 Readers Pay Selective Attention to Task-relevant Articles?

The readers tended to pay selective attention to the task-relevant articles; they more often accessed these articles and spent more time on them. Most of the interview responses confirmed that the participants were thinking about relevance when determining whether to access an article or to continue reading it.

These results are similar to the Spanish L1 research, which the study loosely replicated (Rodicio, 2015). It also corroborates research suggesting Japanese tend to have digital reading competency, at least in their L1 (OECD, 2011). Finally, the findings are also similar to single-document research of Japanese L2 learners of English, who spent relatively more time on relevant sections (Prichard & Atkins, 2019).

Nevertheless, as in previous research, there was room for improvement in terms of the participants' strategic competence. Most clicked on irrelevant pages and many students spent significant time on the first irrelevant page. The interview responses confirmed that a few participants did not always maintain task awareness; several replied they were not thinking or that there was no reason for clicking on specific links and reading the articles. Some participants also did not examine the search page much, especially those who clicked the first irrelevant page. Those that viewed the search page the least at the start of the reading time more often selected an irrelevant article first. This is similar to OECD research (2011), which showed that careful examination of the first page was key.

Another ineffective strategy shown by several participants (and verified by some interview participants) was closely reading irrelevant articles hoping to find some relevant points noted in them. This proved to be detrimental, as is discussed below.

Does Attention to Task-relevant Pages Predict Ability to Recall Relevant Content?

As in previous L1 research (Rodicio, 2015), spending relatively more time on relevant articles made a significant difference on the post-task recall score. Access time on relevant articles was beneficial, and reading time on irrelevant articles was even more detrimental. This impact was similar to previous L2 research involving a single website (Prichard & Atkins, 2019). That fact that time focusing on irrelevant content had strong negative correlation with task performance in these two studies follows literature suggesting that reading irrelevant content may exhaust working memory or interfere with one's ability to effectively store or recall relevant information.

Another strategy that correlated with the recall score was clicking on relevant articles numerous times to read or review the relevant content. This supports OECD findings (2011), though that study showed a ceiling effect.

Like previous research (Rodicio, 2015), the ratio of relevant versus irrelevant articles accessed did not significantly affect post-task performance. This is because, while the number of times accessing relevant sites was a factor, the number of times participants clicked on irrelevant

articles had absolutely no effect. This is somewhat surprising since access *time* on irrelevant articles was highly impactful in a negative way. This suggests that clicking on a link to an irrelevant article was not detrimental as long as participants quickly exited the page. Indeed, some interview participants responded they had clicked irrelevant links by mistake or to check if there were some relevant facts noted, but most responded they soon stopped reading after quickly noticing there was no relevant content. This reflects the metacognitive skill of monitoring task-based comprehension.

How do participants deal with relevant but repeated content?

A secondary purpose of this research was to examine if participants read repeated content already presented in another article or if they simply skimmed or ignored it. This could potentially be a large factor in multi-document reading, yet it has not been frequently discussed in reading research. In this study, Article 5, which repeated content from an earlier article read (Article 3), was viewed by most of the participants, albeit for a shorter time. Similarly, in the two other relevant articles with partially repeated content, the repeated sections were linearly read about three-fourths of the time. Therefore, overall, only a minority of the students skipped over repeated content in this study.

Interview participants replied that they re-read content from different articles to better understand the relevant content or to remember it for the recall task. These reasons seemingly suggest metacognition. However, overall, reading Article 5 actually had a reverse correlation with the recall score. Readers who focused on content that was both relevant and novel performed best. Therefore, at least in this task, it was better to avoid reading repeated content. This warrants further research.

Pedagogical Implications

Researching online in English is an extremely common, real-world task for many learners, yet multi-document task-based reading is often overlooked in the L2 classroom. This study suggested that, although many participants performed adequately, there was room for improvement overall. Readers who demonstrated the most task awareness (those that spent more time reading relevant and novel content) outperformed others. Considering metacognitive strategies did have an effect and some learners lacked competency, it could be concluded that some learners could potentially benefit from instruction.

Experience with task-based reading online in English may be enough to improve participants' performance over time. Any task where learners need to select and read relevant content could build metacognitive competence related to planning and monitoring. However, explicit strategy training may also be necessary for certain learner populations. This could involve instruction, practice, and reflection on task planning, formulating effective search terms, evaluating search results, previewing websites, navigation skills, and maintaining task awareness while reading. Prichard and Atkins (2019) found that training on previewing texts and maintaining task awareness led to significant improvements in performance. However, this only involved single document reading. Empirical research is needed to examine the effect of L2 instruction and strategy training for researching multiple documents.

Since multi-document reading is highly complex, strategy training may not be beneficial for all students; in L1 research, training using multiple-texts did not prove to be helpful for those

with a low-working memory (Naumann, et al., 2008). However, simple training on avoiding irrelevant content could potentially preserve cognitive resources for overly taxed readers. This could be especially beneficial for L2 learners, who may be additionally burdened by having to read texts above their proficiency level.

Limitations and Suggestions for Follow-up Research

This research study had several limitations which should be addressed in follow-up research. First, the number of participants was relatively few. Though having 24 participants is similar to the norm for L2 research involving time consuming methods such as eye tracking (Godfroid, 2019), the study should be replicated before we are confident in the results. Second, note taking was not allowed though it is common in research tasks. Since the participants needed to remember the relevant factors, working memory was a potential factor not controlled for. Finally, the articles were not authentic, as it was deemed necessary to carefully control the text length and difficulty. While this was necessary for empirical research, it should be remembered that the results may have been different had the search page yielded millions of authentic texts.

CONCLUSION

Researching multiple documents online from a search engine page is an extremely common and important activity that has largely been ignored in L2 education and research. Therefore, this study examined the task-awareness and navigation strategies of Japanese English learners, who were provided a search page with both relevant and irrelevant results. The experiment found that many participants did not closely examine the search results and several participants accessed and read irrelevant pages, especially the first one. However, although participants spent a surprisingly long time reading the first article, most participants clicked on and spent more time reading task-relevant pages after that. The study also showed that most participants re-read content already processed on other articles.

The experiment also examined the relationship between reading performance and post-reading recall scores. The results showed that participants who spent more time on relevant articles compared to irrelevant articles performed better, as in previous research. However, one relevant article had repeated content from another article, and reading this text had a reverse correlation with task performance. Overall, participants who focused on articles that were both relevant and novel performed much better on the task. While further research is needed, this further highlights the importance of metacognitive reading skills, such as selective attention to needed content.

Further research is needed among other learner populations to determine if the findings are relevant to them. Moreover, studies examining the effect of task-based reading assignments and strategy training should also be carried out. Such research findings could inform L2 instruction that could potentially improve learners' metacognitive strategy usage and navigation skills, leading to more effective online readers.

ACKNOWLEDGEMENT

This study was supported by Grant-in-Aid for Scientific Research (C) 18K00767 and 19K00662.

REFERENCES

Anmarkrud, Ø., McCrudden, M. T., Bråten, I., & Strømsø, H. I. (2013). Task-oriented reading of multiple documents: Online comprehension processes and offline products. *Instructional Science*, *41*(5), 873-894. https://doi.org/10.1007/s11251-013-9263-8

Bax, S., & Weir, C. J. (2012). Investigating learners' cognitive processes during a computer-based CAE reading test. *Cambridge ESOL, Research Notes, 47*, 3-14. http://www.cambridgeenglish.org/images/22669-rv-research-notes-47.pdf

Bråten, I., & Strømsø, H. I. (2011). Measuring strategic processing when students read multiple texts. *Metacognition and Learning*, *6*(2), 111-130. https://doi.org/10.1007/s11409-011-9075-7

The British National Corpus (2007). *BNC XML Edition*. Distributed by Oxford University Computing Services on behalf of the BNC Consortium.

Cabinet Office, Government of Japan (2014). 平成 25 年度 我が国と諸外国の若者の意識に関する調査.

Carenini, G., Cheung, J. C. K., & Pauls, A. (2013). Multi-document summarization of evaluative text. *Computational Intelligence*, *29*(4), 545-576. https://doi.org/10.1111/j.1467-8640.2012.00417.x

Chang, Y. F. (2006). On the use of the immediate recall task as a measure of second language reading comprehension. *Language Testing, 23*(4), 520-543. https://doi.org/10.1191/0265532206lt340oa

Coiro, J., Knobel, M., Lankshear, C., & Leu, D. J. (2008). *Handbook of research on new literacies*, 1-21. Routledge: New York.

Cress, U., & Knabel, O. B. (2003). Previews in hypertext: Effects on navigation and knowledge acquisition. *Journal of Computer Assisted Learning, 19*(4), 517-527. https://doi.org/10.1046/j.0266-4909.2003.00054.x

Denscombe, M. (2014). *The good research guide: For small-scale social research projects* (5th ed.). London, England: Open University Press.

ETS (2017) Test & Score Data Summary for TOEFL iBT® Tests. www.ets.org/s/toefl/pdf/94227_unlweb.pdf

Fajardo, I., & Villalta, E. (2016). Are really digital natives so good? Relationship between digital skills and digital reading. *Anales de Psicología*, *32*(1), 89.

Godfroid, A. (2019). *Eye Tracking in Second Language Acquisition and Bilingualism: A Research Synthesis and Methodological Guide*. New York: Routledge.

Ghafournia, N., & Afghari, A. (2013). Relationship between reading proficiency, strategic competence, and reading comprehension test performance: A study of Iranian EFL learners. *International Education Studies*, *6*(8), 21-30.

Guo, L. (2018). Modeling the relationship of metacognitive knowledge, L1 reading ability, L2 language proficiency and L2 reading. *Reading in a Foreign Language*, *30*(2), 209-231.

Hirata, Y., & Hirata, Y. (2010). Students' attitudes towards web searching. In P. Tsang, S. K. S. Cheung, V. S. K. Lee, & R. Huang (Eds.) *International Conference on Hybrid Learning* (pp. 127-136). Berlin: Springer. https://doi.org/10.1007/978-3-642-14657-2_13

Karimi, M. N. (2015). L2 multiple-documents comprehension: Exploring the contributions of L1 reading ability and strategic processing. *System, 52*, 14-25. https://doi.org/10.1016/j.system.2015.04.019

Karimi, M. N. (2017). The Mediated/Unmediated Contributions of Language Proficiency and Prior Knowledge to L2 Multiple-Texts Comprehension: A Structural Equation Modelling Analysis. *Applied Linguistics*, *39*(6), 912-932. https://doi.org/10.1093/applin/amw059

Kuiper, E., Volman, M., & Terwel, J. (2005). The Web as an information resource in K–12 education: Strategies for supporting students in searching and processing information. *Review of Educational Research*, *75*(3), 285-328. https://doi.org/10.3102/00346543075003285

Latif, M. M. A. (2019). Eye-tracking in recent L2 learner process research: A review of areas, issues, and methodological approaches. *System*, *83*, 25-35. https://doi.org/10.1016/j.system.2019.02.008

McCrudden, M. T., Magliano, J. P., & Schraw, G. (2011). Relevance in text comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 1-18). Greenwich, CT: Information Age Publishing.

McCrudden, M. T., Stenseth, T., Bråten, I., & Strømsø, H. I. (2016). The effects of topic familiarity, author expertise, and content relevance on Norwegian students' document selection: A mixed methods study. *Journal of Educational Psychology*, *108*(2), 147. https://doi.org/10.1002/rrq.197

Naumann, J., Richter, T., Christmann, U., & Groeben, N. (2008). Working memory capacity and reading skill moderate the effectiveness of strategy training in learning from hypertext. *Learning and Individual Differences*, *18*(2), 197-213.

OECD. (2011). PISA 2009 results: students on line. *In Digital technologies and performance, Vol. VI*, Paris, France: OECD publications.

Peshkam, A., Mensink, M. C., Putnam, A. L., & Rapp, D. N. (2011). Warning readers to avoid irrelevant information: When being vague might be valuable. *Contemporary Educational Psychology*, *36*, 219-231. https://doi.org/10.1016/j.cedpsych.2010.10.006

Phakiti, A. (2003). A closer look at the relationship of cognitive and metacognitive strategy use to EFL reading achievement test performance. *Language Testing, 20*(1), 26-56. https://doi.org/10.1191/0265532203lt243oa

Prichard, C. (2014). Reading strategy use of low- and high-proficiency learners and the effect of reading instruction. Bulletin of Higher Education Okayama University, 10, 115–122.

Prichard, C., & Atkins, A. (2016). Evaluating L2 readers' previewing strategies using eye tracking. The Reading Matrix, 16(2), 110–130.

Prichard, C., & Atkins, A. (2019). Selective attention of L2 learners in task-based reading online. *Reading in a Foreign Language, 31(*2), 269-290.

Readable.io (2018). How readable is your writing? Retrieved from https://readable.io/

Reynolds, R. E., Shepard, C., Lapan, R., Kreek, C., & Goetz, E. T. (1990). Differences in the use of selective attention by more successful and less successful tenth-grade readers. *Journal of Educational Psychology*, *82*, 749-759. https://doi.org/10.1037/0022-0663.82.4.749

Rodicio, H. G. (2015). Students' evaluation strategies in a Web research task: Are they sensitive to relevance and reliability? *Journal of Computing in Higher Education*, *27*(2), 134-157. https://doi.org/10.1007/s12528-015-9098-1

Romly, R., Badusah, J., & Maarof, N. (2017). Metacognitive online reading strategies in reading academic texts among ESL university students. *Proceedings of Seminar on Transdisciplinary Education*, 113-121. Retrieved from https://sted2017.files.wordpress.com/2016/12/1-13-raihana-romly-jamaludin-badusah-and-nooreiny-maarof.pdf

Sakurai, N. (2015). The influence of translation on reading amount, proficiency, and speed in extensive reading. *Reading in a Foreign Language, 27*, 96-112.

Salmerón, L., Cañas, J. J., Kintsch, W., & Fajardo, I. (2005). Reading strategies and hypertext comprehension. *Discourse processes*, *40*(3), 171-191. https://doi.org/10.1207/s15326950dp4003_1

Salmerón, L., Kammerer, Y., & García-Carrión, P. (2013). Searching the Web for conflicting topics: Page and user factors. *Computers in Human Behavior*, *29*(6), 2161-2171. https://doi.org/10.1016/j.chb.2013.04.034

Sheorey, R., & Mokhtari, K. (2001). Differences in the metacognitive awareness of reading strategies among native and nonnative readers. *System, 29*, 431–449. https://doi.org/10.1016/S0346-251X(01)00039-2

Smallwood, J., Fishman, D. J., & Schooler, J. W. (2007). Counting the cost of an absent mind: Mind wandering as an underrecognized influence on educational performance. *Psychonomic Bulletin & Review*, *14*, 230-236. https://doi.org/10.3758/BF03194057

Stevens, C., & Bavelier, D. (2012). The role of selective attention on academic foundations: a cognitive neuroscience perspective. *Developmental Cognitive Neuroscience, 2*(*Suppl.* 1), S30-S48. https://doi.org/10.1016/j.dcn.2011.11.001

Tipper, S. P., & Baylis, G. C. (1987). Individual differences in selective attention: The relation of priming and interference to cognitive failure. *Personality and Individual Differences, 8*, 667-675. https://doi.org/10.1016/0191-8869(87)90064-X

White, H., McConnell, E., Clipp, E., Branch, L. G., Sloane, R., Pieper, C., & Box, T. L. (2002). A randomized controlled trial of the psychosocial impact of providing internet training and access to older adults. *Aging & Mental Health*, *6*, 213–221. https://doi.org/10.1080/13607860220142422

Zhang, L. J. (2010). A dynamic metacognitive systems account of Chinese university students' knowledge about EFL reading. *TESOL Quarterly, 44*, 320–353. https://doi.org/10.5054/tq.2010.223352

Zhang, L. J., & Wu, A. (2009). Chinese senior high school EFL students' metacognitive awareness and reading-strategy use. *Reading in a Foreign Language*, *21*, 37-59.

Zhang, L., & Seepho, S. (2013). Metacognitive Strategy Use and Academic Reading Achievement: Insights from a Chinese Context. *Electronic Journal of Foreign Language Teaching, 10*(1), 54-69.

Caleb Prichard *is an Associate Professor at Okayama University in Japan. His main research interests are reading strategy use, humor competency training, and language program administration.* Email: prichard@okayama-u.ac.jp

Andrew Atkins *is an Associate Professor at Kindai University in Osaka, Japan. His research interests include reading fluency development, intercultural communication, and testing.*
Email: andrew@kindai.ac.jp