Effect of memory load on eye movement control: A study using the reading span test

Miyuki Azuma
Osaka University

Takehiro Minamoto
Osaka University

Ken Yaoi
Kyoto University

Mariko Osaka
Osaka University

Naoyuki Osaka
Kyoto University

We investigated the effect of memory load on attentional control using the Reading Span Test (RST), a task that requires working memory capacity. Previous studies have shown that a shortage of working memory resources leads to a deficit of inhibition of task-irrelevant information and that memory load affects eye movement control. Here, we recorded eye movement and integrated it with RST performance. Total fixation time and the number of regressions showed a memory load effect with the to-be-remembered word, and RST performance was also affected under high memory load. We concluded that a shortage of working memory resources caused by memory load prevents flexible eye movement control and may cause a deficit in inhibitory control based on intrusion errors.

Keywords: working memory, reading span test (RST), memory load, intrusion error, total fixation time, regression

Introduction

Studies of eye movement during reading have a long history dating back to the late 19th century, and several models of eye movement control in reading have been proposed (Rayner, 2008). In everyday life, we strongly rely on visual attentional control for higher cognitive activities like reading. To understand the meaning of text, precisely-controlled eye movements that provide enough fixation time to read the text are used (Osaka & Osaka, 1994). Furthermore, an individual utilizes attentional control in the brain to process and store information in order to understand what was just read (Azuma et al., 2012).

Working memory supports the simultaneous processing and storing of information by controlling attention (Baddeley, 1986; Baddeley, 2007). It has a particularly critical role in comprehension processes during text reading (Daneman & Carpenter, 1980), which require the decoding of incoming text information, recognizing the information, storing it for short periods until it is integrated into a sentence, and finally retaining them for integration with subsequent incoming information. Working memory deals with such processes simultaneously to comprehend texts fluently.

Because working memory resources are limited, resources must be allocated properly for cognitive tasks, as otherwise the simultaneous storage and processing of information will fail. Focusing on the limited capacity of working memory, some complex span tests were invented. One of the most frequently employed working memory tasks is the Reading Span Test (RST), the first of which was developed by Daneman & Carpenter (1980) in English. In the RST, participants are asked to read a series of short sentences orally and simultaneously memorize the last word of each sentence. After reading a certain number of sentences, participants are asked to recall the last word of each sentence in a specific order. That is to say, participants are asked to perform a dual task: understand a sentence by reading aloud and memorize the last word of each sentence. Scores in the RST have shown higher correlation with reading comprehension than scores from a more traditional short-term memory task, indicating the
RST is a reliable measure of individual working memory capacity (Just & Carpenter, 1992; Daneman & Merikle, 1996; Osaka & Osaka, 1992; Osaka, Osaka & Groner, 1993). Additionally, it is assumed that the RST measures an individual’s efficiency at comprehending text, that is, the more effectively an individual can process sentences, the more resources the individual has for storing the words.

Kane, Conway, Hambrick, and Engle (2007) suggested that working memory span tasks primarily reflect domain general executive processes and secondarily domain specific processes. Executive processes help maintain access to to-be-remembered items while a subject is conducting a processing task or distracted by irrelevant information in working memory span tasks. In fact, many studies have shown that individual working memory span differences are reflected in various attention tasks, such as the dichotic listening task, visual search task, antisaccade task, and stroop task (Colflesh & Conway, 2007; Conway, Cowan, & Bunting, 2001; Fukuda, & Vogel, 2011; Kane, Bleckley, Conway, & Engle, 2001; Kane, & Engle, 2003; Vogel, McCollough, & Machizawa, 2005). The results from these studies support the theory that individuals with high working memory span have flexible and effective attentional control. On the other hand, those with low working memory span are poor at ignoring irrelevant information and concentrating their attention on important information (Borella, Ludwig, Fagot, & De Ribau-pierre, 2011; Carretti, Cornoldi, De Beni, & Romanò, 2005; Chiappe, Hasher, & Siegel, 2000; De Beni, Palladino, Pazzaglia, & Cornoldi, 1998; Gernsbacher, 1993; Osaka, Nishizaki, Komori, & Osaka, 2002).

Focusing on attentional control in the RST, previous studies have shown that individuals with low working memory span tend to mislead their attention to information that is not required in the primary goal. One typical and resulting error is the intrusion error, which describes a participant answering a word in the sentence other than the to-be-remembered word. The study of Chiappe et al. (2000) showed that older adults, who generally have lower working memory spans than younger adults, have higher intrusion errors. The authors assumed that this type of error reflects a failure in preventing goal-irrelevant information from entering working memory. Osaka et al. (2002) showed that the low RST performers were often distracted by a word important for comprehending the sentence. Later experiments by the group using fMRI found different brain activities between high and low span individuals during the period of reading the RST sentence and recognizing the to-be-remembered word (Osaka, 2006; Osaka, Komori, Morishita, & Osaka, 2007; Osaka, & Osaka, 2007).

Previous Studies about Eye Movement using RST

Attention control in working memory can also be investigated by measuring eye movements. Studies have shown that individuals with high working memory capacity can utilize prior knowledge and effectively control their eyes to comprehend text (Kaakinen, Hyönä, & Keenan, 2003), with high span individuals showing better text integration under smaller visual fields, a condition where working memory resources are in greater demand (Osaka, & Osaka, 2002). This property suggests that high span individuals have flexible eye movement control for demanding tasks.

We previously measured eye movements using the Japanese RST (Azuma, Ikeda, Minamoto, Osaka, & Osaka, 2012), which was first introduced by Osaka & Osaka (1994). Some differences in the English and Japanese version include the position of the to-be-remembered word and the role of the to-be-remembered word in the sentence. The last word in an English sentence is quite variable, as it can be a noun, verb, adjective, or adverb, but is also an important word for understanding the sentence. On the other hand, Japanese sentences almost always end with a verb. To account for this difference, the to-be-remembered word in the Japanese RST is underscored even if it is not the most important word of the sentence. This last point means that performers of the Japanese RST must continue to read aloud after they read the to-be-remembered word, that is, the other words in the sentence act as distractors that participants must resist remembering. In a previous RST study, we implemented two conditions, easy and difficult (two- and five-sentences), and compared high and low working memory capacity groups. A difference between groups was only seen in the difficult condition, with the high working memory capacity group spending longer total fixation time on the to-be-remembered word than the low working memory capacity group. This result indicated that individuals with low working memory capacity cannot control their attention efficiently when their memory capacities are overloaded. We concluded that the shortage of working memory resources caused by the heavy memory
load leads to poor attentional control and less performance in the RST.

Studies using the RST to investigate the relationship between memory load and eye movements have reported that as memory load becomes higher, participants tend to gaze the to-be-remembered word longer. Carpenter & Just (1989) used the English version of the RST (i.e., participants have to remember the final word of the sentence) and found that high RST performers read sentences faster and spend time on the to-be-remembered word longer when required to keep to-be-remembered words from the previous sentence. Kaakinen & Hyönpää (2007) used the Finnish version of the RST, in which the to-be-remembered word is also the final word of the sentence, and found that participants spent less time on the first and middle words of a sentence and more time on the to-be-remembered word as memory load increased. Engle, Cantor, & Carullo (1992), using an English version of the RST in which the to-be-remembered word is separate and semantically unrelated to the sentence, found the high-span group spends more time on the first, last and to-be-remembered words of the sentence than the low-span group as memory load increased. Further, the high-span group kept their viewing time constant under memory load, whereas the low-span group showed a peak viewing time at a certain memory load that decreased with greater load. These studies used the RST where the position of the to-be-remembered word is fixed, which simplifies the task. There is greater difficulty in the Japanese RST, because the variable position of the to-be-remembered word results in the subsequent words of the sentence acting as distractors.

The purpose of the present study was to investigate whether working memory shortage leads to poor eye movement control in the Japanese RST. We investigated the relationship between working memory resources and attentional control by comparing eye movements on to-be-remembered and goal-irrelevant words in the Japanese version of RST, which is considered to require more attentional control than other versions, under different memory loads.

Method

Participants

Participants were undergraduate and graduate students at Kyoto University (N = 32, 20 males and 12 females). The mean age was 21.94 years (SD = 2.40, age range 20 - 31). Fourteen participants were removed from the data analysis because of technical issues: repetitive failure of calibration (N = 12) or extreme head movements (N = 2). All participants reported normal or corrected-to-normal visual acuity and gave their informed consent in accordance with the Department of Medicine, Kyoto University.

Eye Movement Recording

Eye movements were recorded with a Tobii T120 eye tracker (Tobii Technology AB, Danderyd, Sweden) using a 120 Hz data sampling rate and an automatic calibration procedure that included five calibration points: one center and four corners. The maximum error in the calibration was 30 pixels in diameter. A chin rest was used to restrict head movements at a viewing distance of 57 cm. Because the eye tracking system had a large freedom of head movement (30 x 22 x 30 cm), we could collect good-quality data under oral reading.

Materials

The Japanese RST was conducted prior to the experiments in order to measure individual working memory capacity. We calculated the individual span score following the procedure showed in Osaka & Osaka (1994). The participants were told to recall the words in serial order and not to report the last sentence to-be-remembered word first. The mean span score of all participants was 2.83 (span range 2.0 - 4.5) and of the 18 participants whose eye movement data were analyzed was 2.66 (span range 2.0 – 4.0).

Another version of the Japanese RST was employed for eye movement experiments. Target words were underlined in red. All sentences consisted of Japanese Kana and Kanji and had a mean length of 25.52 characters (24-27 characters) and 31.47 moras (28-34 moras). The mean sentence length of the first sentence was 25.33 characters and 31.38 moras; second sentence, 25.92 characters and 31.62 moras; third sentence, 25.50 characters and 31.63 moras; and fourth sentence, 25.33 characters and 31.25 moras. We selected one additional word (comparative word) from each sentence in order to compare eye
movement patterns for the to-be-remembered word. The to-be-remembered and comparative words were all nouns consisting of two Kanji characters. The mean word mora was 3.65 for the to-be-remembered word (3-4 moras, mean mora are 3.48 for the first sentence, 3.50 for the second sentence, 3.29 for the third sentence, and 3.86 for the fourth sentence) and 3.50 for the comparative word (2-4 moras, mean mora were 3.39 for the first sentence, 3.57 for the second sentence, 3.03 for the third sentence, and 3.59 for the fourth sentence). The size of the letters was 34 pixels and extended a visual angle of 0.9 degrees. The visual angle of each sentence was approximately 24 degrees.

Procedure

Participants were asked to read a sentence and memorize the to-be-remembered word simultaneously. The task consisted of twelve trials all using the four-sentence condition (Figure 1).

![Figure 1. Example of a presentation sequence.](image)

English translation of the first sentence is “I expect the discussion about this unreasonable rule to begin”. The to-be-remembered word is “rule”. The second sentence translates to, “I wonder why so many people are excited about this competition”, and the to-be-remembered word is “competition”. A red circle was presented before each sentence and acts as a fixation point. Participants were instructed to remember the underlined word while reading aloud the whole sentence. A line of asterisks was presented after the final sentence and acted as a cue.

Before each sentence was presented, participants were instructed to gaze at a fixation point presented for 1,000 ms. The fixation point was replaced by the first letter of a sentence, following the procedure of Azuma et al. (2012), in order to make the participants fixate there at the beginning of every sentence. The entire sentence and the cue underline of the to-be-remembered word was presented simultaneously and remained on the monitor until the next fixation point appeared. Participants were required to read each sentence aloud as soon as it was presented.

We employed the experimenter-paced method. When participants finished reading the sentence orally, the presentation of the sentence was terminated by the experimenter, and the fixation point immediately appeared. After all four sentences in a trial were presented, a line of asterisks appeared, which cued the participants to report the to-be-remembered words orally. The participants were told to recall the words in serial order. Each experiment lasted approximately 15 minutes for the twelve trials.

Data Analysis

Eye movement data were analyzed using Tobii Studio 2.3.2. When the eye remained more than 50 ms in the word area, which was set as a rectangle surrounding the to-be-remembered and comparative words, we defined it as a fixation using an automatic fixation detection algorithm (Tobii Fixation Filter, detection radius of 20 pixels on a screen with a resolution of 1280 x 1024 pixels). The RST score is the total number of correctly recalled words. In the case that a word was reported in the wrong order, it was included in the correctly recalled word in correct order (i.e., when the second to-be-remembered word was reported third, it was counted as a correctly recalled word for the second sentence). Data were statistically analyzed using STATISTICA (StatSoft JAPAN, Tokyo, Japan).

Results

RST Score

The mean number of correctly recalled words. Table 1 shows the mean number of correctly recalled words. One-way ANOVA (sentence order: first, second, third, and fourth) showed a main effect of sentence order, \(F(3, 51) = 13.42, \ p < .01, \ \eta^2_p = 0.44\). Tukey’s HSD analysis showed that the word in the third sentence was less recalled than in the other sentences (all \(p < .01\)), and that the recalled word in the fourth sentence was more recalled (\(p < .05\) compared with the first sentence, and \(p < .01\) compared with the second and third sentences).
Intrusion error. Table 1 shows the mean number of intrusion errors for each sentence. Intrusion error is the error when a participant recalls a word different from the to-be-remembered word in the current trial. In the case that an intrusion word was reported in the wrong order, it was counted as error in the correct order (i.e., when the non-to-be-remembered word in the second sentence was reported third, it was counted as an intrusion error of second sentence). One-way ANOVA (sentence order: first, second, third, and fourth) showed significant effect of memory load, $F(3, 51) = 6.09, p < .01, \eta^2_p = 0.26$. Tukey’s HSD analysis showed that participants made more intrusion error in the fourth sentence than the others ($p < .01$ compared with the first and second sentences, and $p < .05$ compared with the third sentence).

Reading Time

The mean sentence reading time was (in order of first to fourth sentence) 5.184 ms ($SD = 493.3$), 5.169 ms ($SD = 623.7$), 5.338 ms ($SD = 594.3$), and 5.168 ms ($SD = 648.5$). One-way ANOVA was conducted and found a main effect of the sentence order, $F(3, 51) = 4.96, p < .01, \eta^2_p = 0.23$. Tukey’s HSD analysis showed that the third sentence required a longer reading time (all $p < .05$).

Eye Movement Results

Eye movement results consisted of two indicators: total fixation time, which is a reflection of the total amount of attention, and the number of regressions, which reflects additional attention.

Total fixation time. Total fixation time was defined as the sum of the time for all fixations within the to-be-remembered or comparative word areas. The mean total fixation time for each sentence order is shown in Table 2.

We conducted a mixed 2 (word type: to-be-remembered word, comparative word) x 4 (sentence order: first, second, third, and fourth sentence) ANOVA. Both word type and sentence order were within-participant factor. The effect of word type was significant, $F(1, 17) = 19.6, p < .01, \eta^2_p = 0.54$, as too was the effect of sentence order, $F(3, 51) = 4.66, p < .01, \eta^2_p = 0.22$. These results indicate that eye gazing was longer on the to-be-remembered word than the comparative word, and it was longer in the first and second sentences than in the fourth. Moreover, since the interaction nearly showed tendency, $F(3, 51) = 2.14, p = .11, \eta^2_p = 0.11$, we analyzed each word type separately to examine the effect of memory load in more detail.

The result of one-way (sentence order: first, second, third, and fourth sentence) ANOVA for the to-be-remembered word showed a significant main effect of the sentence order, $F(3, 51) = 4.49, p < .01, \eta^2_p = 0.21$. Tukey’s HSD analysis showed that participants gazed at the to-be-remembered word in the fourth sentence for a shorter period than in the first or second sentence ($p < .01$ for the first sentence, $p < .05$ for the second sentence). The same one-way ANOVA was conducted for the comparative word, but the effect of the sentence order was not significant, $F(3, 51) = 0.84, ns, \eta^2_p = 0.05$. In summary, memory load decreased the time spent on the to-be-remembered word but had no effect on the comparative word.

The number of regressions. The number of regressions is the total number of eye movements in a word area minus the first entry. When a word was skipped (not fixated), the trial was excluded from the analysis. The skip rate was 3% for the to-be-remembered word and 9% for the comparative word. In the case that the eyes moved

<table>
<thead>
<tr>
<th>Load</th>
<th>Correctly Recalled Words</th>
<th>Intrusion Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Sentence 1</td>
<td>8.22</td>
<td>1.59</td>
</tr>
<tr>
<td>Sentence 2</td>
<td>7.56</td>
<td>2.20</td>
</tr>
<tr>
<td>Sentence 3</td>
<td>6.11</td>
<td>2.14</td>
</tr>
<tr>
<td>Sentence 4</td>
<td>9.33</td>
<td>1.50</td>
</tr>
</tbody>
</table>
into a word area only once, the number of regressions was defined as zero. The mean number of regressions for each sentence order is shown in Table 2.

We conducted a mixed 2 (word type: to-be-remembered word, comparative word) x 4 (sentence order: first, second, third, and fourth sentence) ANOVA. Both word type and sentence order were within-participant factor. The effect of word type was significant, $F(1, 17) = 114.89, p < .01, \eta^2_p = 0.87$, indicating that the to-be-remembered word had more regression than the comparative word. Similarly, the effect of memory load was significant, $F(3, 51) = 5.27, p < .01, \eta^2_p = 0.24$. According to Tukey’s HSD analysis, the to-be-remembered word in the first and second sentences was more regressed than those in the first sentence ($p < .01$). The interaction between word type and sentence order was also significant, $F(3, 51) = 2.89, p < .05, \eta^2_p = 0.15$. Therefore, we examined the effect of memory load on the to-be-remembered word and the comparative word separately.

One-way (sentence order: first, second, third, and fourth sentence) ANOVA found the to-be-remembered word had a significant main effect of memory load, $F(3, 51) = 5.27, p < .01, \eta^2_p = 0.24$. According to Tukey’s HSD analysis, the to-be-remembered word in the first and second sentences was more regressed than in the fourth sentence ($p < .01$ and $p < .05$, respectively). However, the comparative word showed no significant effect of memory load, $F(3, 51) = 0.82, ns, \eta^2_p = 0.05$. Thus, the number of regressions indicates that participants had difficulty controlling their attention when gazing at the to-be-remembered word.

## Table 2

<table>
<thead>
<tr>
<th>Load</th>
<th>Total Fixation Time (ms)</th>
<th>Number of Regressions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target Word</td>
<td>Comparative Word</td>
<td>Target Word</td>
</tr>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Sentence 1</td>
<td>642.27</td>
<td>213.47</td>
<td>384.45</td>
</tr>
<tr>
<td>Sentence 2</td>
<td>631.00</td>
<td>236.64</td>
<td>400.15</td>
</tr>
<tr>
<td>Sentence 3</td>
<td>591.64</td>
<td>214.51</td>
<td>405.39</td>
</tr>
<tr>
<td>Sentence 4</td>
<td>526.06</td>
<td>255.00</td>
<td>370.80</td>
</tr>
</tbody>
</table>

### Discussion

#### Memory Load and Attentional Control

In the present study, we found a significant effect of memory load on the to-be-remembered word in the mean number of correctly recalled words, the total fixation time, the number of regression, supporting the theory that a shortage of working memory resources leads to poor attentional control including eye movement control and inhibitory control. Participants had sufficient working memory resources when reading the first and second sentences, as they could spend more time and frequency on task-relevant information. However, by the fourth sentence, their working memory resource was consumed by storage and they lacked efficient attentional control. In the third sentence, when the participants’ capacity reached its limit, memory effect was reflected by longer reading times. Despite the longer reading time, the total fixation time and the number of regressions did not increase. We suspect that the participants had difficulty in whole sentence comprehension, suffering from conflict between word maintenance and sentence processing. Because the to-be-remembered word was in the middle of the sentence, participants encountered it before completing whole sentence comprehension. Osaka (2002) claimed that the RST requires an inhibition operation at multiple phases during the task and that this inhibition operation should be controlled carefully. Each word is important for understanding a sentence, and the to-be-remembered word is crucial for task achievement. RST performers should maintain activation of the to-be-remembered word appropriately so as not to interfere with whole sentence comprehension but also not to forget the to-be-remembered word completely. The difficulty in this task decreased the number of correctly recalled words in the third sentence and slightly decreased the
total fixation time and the number of regressions.

The memory load effect was also significant in the
fourth sentence based on decreased eye movements. Al-
though the RST score was best for the fourth sentence, the
intrusion error was also highest. One explanation for this
result is that the participants relied too much on the re-
cency effect, speeding the sentence reading time to re-
duce the storing duration. Such a strategy risks insuffi-
cient attention for crucial information. In our previous
study (Azuma et al., 2012), we showed that high RST
performers spend more time on task-relevant information
under a high memory condition, which suggests they
maintain flexible eye movement even at later sentences in
a trial. Robert, Borella, Fagot, Lecerf, & De Ribaupierre
(2009) found intrusion errors indicate a failure in inhibi-
tory control due to the availability of fewer resources.
The result of the present study supports it, and moreover
we suggest that the inefficient eye movement control
leads to more intrusion errors.

Comparison with Other RSTs

We show memory load has an effect on total fixation
time and the number of regressions, which does not agree
with previous studies where participants had faster read-
ting times of sentences and longer fixation times on to-be-
remembered words as memory load increased (Carpenter
& Just, 1989; Kaakinen & Hyönä, 2007). The different
observations may be due to the language of the RST. The
position of the to-be-remembered word in the Japanese
RST is not fixed, which could lead to deficient eye
movement control on the to-be-remembered word. Fur-
ther, the random location requires participants continue to
read aloud after they read the to-be-remembered word
(i.e., participants resist accessing and retaining other
words). Also, to-be-remembered words do not always
have a focus effect, whereas in English and Finnish RSTs
they tend to be an important word in the sentence. These
differences in RSTs result in more attentional control and
more occupancy of working memory resources. Concern-
ing the result of reading time, if we exclude the fourth
sentence from the analysis, the observed reading times
agree with those reported by Friedman & Miyake (2004),
who found that the reading time increased with memory
load under the experimenter-administered condition. The
features of the Japanese RST may also cause this differ-
ence. The version of RST used in their study was similar
to the original version (Daneman & Carpenter, 1980),
where the to-be-remembered word was the sentence-final
word. Thus, subjects could have advantages in both inter-
nal and external attentional control. On the other hand,
the participants in the present study would find it difficult
to simultaneously keep the to-be-remembered word of the
previous sentence and read the present sentence while
reading the third sentence, as mentioned above. The re-
coversy of the reading time in the fourth sentence could
reflect a strategy to shorten the retention time of the to-
be-remembered words. Such a strategy resembles but is
not identical to those used in the English and Finnish
RSTs, where participants read faster in the sentence por-
tion and spend more time on to-be-remembered words. In
our results, as memory load increased, total fixation time
and the number of regressions of the to-be-remembered
word decreased. In other words, participants not only
read the sentence faster, but the to-be-remembered word
too. Predicting the position of the to-be-remembered
word makes it easier to lead attention to task-relevant
information and also to screen out task-irrelevant infor-
mation. Our result is similar to low-span subjects attend-
ing to-be-remembered words whose positions were fixed
(Engle et al., 1992). That study required participants re-
member the content of the sentence for a true-false ques-
tion, which resembles the distractors used in the Japanese
RST. The pattern of eye movement in ours and that study
differed, however, as the peak of the low-span group was
at the middle sentence of the trial, while the peak of the
present study was in the first sentence. Based on the
number of regressions and intrusion errors in the present
study, participants, when having enough working
memory resources, could regress to the to-be-
remembered words and keep task-relevant information in
the focus of attention. However, under the condition that
the participants consumed their working memory re-
sources, they lost flexible eye movements and failed to
keep access to the to-be-remembered word and allowed
task-irrelevant words to be accessed.

Conclusions

We found participants conducting the Japanese RST
need higher attentional control than RST results reported
in other languages. The frequency of fixation on task-
relevant information had a downward tendency with
longer reading time and more intrusion error, which re-
sulted from increased memory load: participants could
control their attention more effectively under lower memory load, but performed poorly at higher memory load. Such eye movements are different from studies using other versions of the RST. The results of intrusion error support the idea that a shortage of working memory capacity leads to deficit inhibitory control. Thus, less available working memory resources caused defective eye movement control and poorer inhibitory control in working memory.

Acknowledgements

This work was presented at the Special Symposium in Honor of Rudolf Groner, 17th European Conference on Eye Movements (ECEM2013) held in Lund, Sweden and supported by grants from the Japan Society for the Promotion of Science (#22220003 to N.O. and (#23240036) to M.O.

References


