THE EFFECT OF WORKING MEMORY CAPACITY AND MENTAL EFFORT ON MONITORING ACCURACY IN TEXT COMPREHENSION

Kenji IKEDA¹² and Shinji KITAGAMI¹

¹Nagoya University
²Japan Society for the Promotion of Science

We investigated the effect of working memory capacity and mental effort on the ability to monitor the accuracy of text comprehension. Participants completed the operation span test to measure working memory capacity (WMC) and read five scientific expository texts. In low-effort condition, participants read intact texts, whereas in high-effort condition, participants read texts that had some letters deleted. After reading, participants assigned comprehension rating to each text and then completed a comprehension test. The result showed a significant interaction between WMC and mental effort. Increased effort improved the accuracy of low-WMC readers but decreased the accuracy of high-WMC readers. Further, in low-effort condition, high-WMC readers monitored their comprehension more accurately than did low-WMC readers. In high-effort condition, low-WMC readers monitored their comprehension more accurately than did high-WMC readers. Our findings showed that both WMC and mental effort affected the ability of participants to monitor their accuracy. These findings are discussed.

Key words: Monitoring accuracy, Working memory, Mental effort

During learning, individuals regulate their process to achieve better performance. For example, people might devote more time to less well-understood material than to material that has previously been well absorbed, and they might end their study efforts when they believe learning has been completed. These self-regulating mechanisms are based on judgments about whether materials have been sufficiently learned. In this way, metacognitive monitoring is important for self-regulated learning (e.g., Dunlosky & Thiede, 2004; Koriat & Goldsmith, 1996; Thiede & Dunlosky, 1999). It is suggested that monitoring plays an important role in self-regulated text comprehension. That is, accurate monitoring leads to more efficient studying and better performance (e.g., Thiede, 1999; Thiede & Anderson, 2003; Thiede, Anderson, & Therriault, 2003).

However, many previous studies of monitoring of comprehension have reported that readers cannot accurately monitor their own comprehension. The correlation between metacomprehension judgments and test performance was only .27 (Maki, 1998). However, metacognitive monitoring plays an important role in self-regulated learning, and thus accurate monitoring leads to more efficient studying and better performance (e.g., Thiede, 1999; Thiede & Anderson, 2003; Thied et al., 2003). Thus, it is important to

Correspondence concerning this article should be addressed to Kenji Ikeda, Department of Psychology, Graduate School of Environmental Studies, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8601, Japan (e-mail: ikeda.kenji@f.mbox.nagoya-u.ac.jp).
investigate the factors that contribute to monitoring process and that can improve its accuracy.

Previous studies suggested that readers infer their comprehension level on the basis of cues (e.g., Anderson & Thiede, 2008; Dunlosky, Rawson, & Middeleton, 2005; Thiede, Griffin, Wiley, & Anderson, 2010). To obtain these cues, readers have to monitor their process during reading. Thus, one factor affecting monitoring is working memory capacity (WMC), because readers need to have sufficient available resources to allocate to monitoring (Dunlosky & Rawson, 2005). Griffin, Wiley, and Thiede (2008) investigated the relationship between WMC and monitoring accuracy in terms of the dual-processing assumption. The model of metacognition assumes object-level and meta-level processing (e.g., Nelson & Narens, 1990, 1994). Object-level processing focuses on text, whereas meta-level processing focuses on mental processes and the text representations. Thus, accurate monitoring requires the allocation of attentional resources to both object- and meta-level processes.

In addition, given that processing text is a primary process, whereas allocation of resources to the meta-level is a secondary process, readers are required to have good attentional control ability. Given that WMC is related to attentional control (e.g., Conway & Engle, 1994; Conway et al., 2005; Kane, Bleckley, Conway, & Engle, 2001), WMC might affect monitoring accuracy. Griffin et al. (2008) investigated this assumption and demonstrated that higher-WMC was associated with more accurate monitoring.

Furthermore, Griffin et al. (2008) also demonstrated a rereading effect. Rereading reduces resources allocated at the object-level; as a result, more resources can be allocated to the meta-level. Thus, in rereading, lower-WMC readers allocate more resources to the meta-level, and their monitoring accuracy improves. These results are consistent with the dual-processing account. Therefore, to obtain the cues needed to infer their own comprehension level, readers need to allocate attention to the meta-level.

On the other hand, it is possible that the increased processing, which leads to the requirement of allocation of more resources to the object-level, also improves monitoring accuracy. According to the increased processing hypothesis, more effortful processing leads to more accurate monitoring (Linderholm & Zhao, 2008). Maki, Foley, Kajer, Thompson, and Willert (1990) investigated whether increased processing effort following the deletion of letters made improvement of accuracy. McDaniel (1984) suggested that deleted-letters text was more difficult to understand than intact text, because deleted-letters text took a longer time to read than intact text. Maki et al. (1990) demonstrated that this manipulation led to improved monitoring accuracy. This improvement of accuracy may be due to a comprehension rating based on ease-of-processing rather than on text retrievability (Rawson & Dunlosky, 2002).

If the effort to processing texts were increased, readers would be unable to allocate sufficient resources to the object-level and, as a result, could not obtain cues needed to infer their own comprehension level. In this situation, given that individuals rely on cues that are available and subjectively salient (Koriat & Ma’ayan, 2005), readers might judge their comprehension level on the basis of ease-of-processing. Such cues might facilitate inferences about comprehension under the assumption that understanding difficult texts
would be difficult, because disruptions in reading fluency serve as important cues to improve accuracy (Miesner & Maki, 2007).

Given that effortful processing is linked to accurate monitoring, it is possible that low-WMC readers may monitor their comprehension more accurately than high-WMC readers do, because low-WMC readers read texts more slowly, and thus with greater effort, than do high-WMC readers (Linderholm & Zhao, 2008). In fact, the bias between prediction and actual performance in low-WMC readers was lower than that in high-WMC readers (Linderholm, Cong, & Zhao, 2008; Linderholm & Zhao, 2008). Thus, low-WMC readers monitored their comprehension more accurately than did high-WMC readers.

However, this assumption based on increased processing hypothesis is inconsistent with dual-processing assumption. In this regard, it is possible that effortful processing may have minimal effect in Griffin et al. (2008). This study found no effect of rereading on performance, although rereading improved performance in other research (Rawson, Dunlosky, & Thiede, 2000). Thus, it would seem that readers constructed a situation model reflecting deeper levels of text representation (Kintsch, 1998; van Dijk & Kintsch, 1983) while reading only once. Hence, it seems that readers processed texts easily and allocated less effort to reading. Griffin et al. (2008) suggests an alternative view, namely, that effortful processing may have minimal effect. Thus, it is possible that the advantage of low-WMC readers in monitoring obtains in conditions where more effort is required to process texts. Considering this possibility, low-WMC readers may monitor their comprehension more accurately than did high-WMC readers.

In the present study, we investigated how WMC and mental effort affected monitoring accuracy\(^1\). To investigate this proposal, we used the deleted-letters paradigm described by Maki et al. (1990). That is, we used intact text in the low-effort condition and deleted-letters text in the high-effort condition. In the low-effort condition, the monitoring of high-WMC readers might be more accurate than that of low-WMC readers, because high-WMC readers would be able to allocate sufficient resources to both levels, whereas low-WMC readers would allocate more resources to processing the text, yielding insufficient resources at the meta-level. In contrast, in the high-effort condition, we hypothesized that the accuracy of low-WMC readers would be higher than that of high-WMC readers. Low-WMC readers may be unable to allocate sufficient resources to object-level; as a result, they may make a judgment on the basis of difficulty. In contrast, high-WMC readers would be unable to allocate sufficient resources to the meta-level.

---

\(^1\) In the present study, we focused on whether readers can discriminate comprehension levels of each text, because this relative accuracy links to self-regulated learning (Thiede, 1999). Thiede et al. (2003) demonstrated that readers, who monitor their comprehension accurately, selected texts rated as lower comprehension rating for restudy phase, and as a result, their comprehension test performance was improved dramatically. Therefore, it is important for self-regulated learning to discriminate more and less comprehended texts, and thus, we focused on relative accuracy in the present study.
METHOD

Participants and Design
Seventy-eight undergraduate and graduate students from Nagoya University participated in the experiment. All participants were native Japanese speakers. A two-way between-subjects design was used. Participants were randomly assigned to a low-effort or a high-effort condition. Participants were categorized as low- or high-WMC readers on the basis of the operation span test (OST), originally developed by Turner and Engle (1989).

Materials
OST. We used the OST to measure WMC. In this task, operation-word strings were presented at the center of a computer monitor [e.g., “(7 × 1) + 6 = 13”, and participants were asked to read aloud, solve, and indicate whether operations were correct maru (yes) or incorrect batsu (no). After answering, participants read words aloud. A new operation-word string was presented immediately after each word was articulated unless three question marks (???) appeared. If three question marks appeared, participants had to recall the words previously presented. Following Conway et al. (2005), the number of operation-word strings presented varied from two to six, with three trials for each set size. The order of the operation-word strings and set sizes were randomized. We used the same operations as Cantor and Engle (1993) and selected words with concreteness and meaningfulness scores of more than 4.0 and ease-of-learning scores of more than 3.5 from among those used by Ogawa and Inamura (1974).

Texts. We used five expository texts (e.g., about permanent magnets or fuel cells) selected from the Japanese science magazine NEWTON; texts were appropriately 200-300 words in length. We used both intact texts and deleted-letter texts. We deleted letters from texts according to Maki et al. (1990). First, texts were divided into idea units (IUs), which are syntactic units, according to the method used by Muramoto (1992). The mean of number of IUs was 27. Next, we deleted letters from words (noun, verbs, or adjectives) after every third IU. Two letters were deleted from words with 4–5 letters, three were deleted from words with 6–7 letters, and four were deleted from words with eight or more letters. When an IU included four or more words but no deletions, we deleted letters in the previous or subsequent IU. Unlike the procedure followed by Maki et al., we underlined words with deleted letters so participants could clearly distinguish them.

Comprehension Tests. Comprehension was tested with two multiple-choice questions for each text. These questions required inference to answer correctly.

Procedure
Participants first completed the OST to measure WMC. Next, participants read all texts. The order of text presentation was randomized, and participants were instructed to understand the texts. Participants read all texts, with each presented one sentence at a time on a computer monitor. Participants pressed the enter key to signal they had read each sentence, prompting the presentation of the next sentence. In the low-effort condition, participants read intact texts, and in the high-effort condition, participants read deleted-letter texts. Participants read at their own paces but were not allowed to review previous sentences. After reading all texts, participants rated their comprehension of each in response to the presentation of each title. Participants were asked to answer the question “How well do you think you understood the text?” on a 7-point scale, ranging from 1 (“very poorly”) to 7 (“very well”). Finally, participants completed comprehension tests. The order of administration for the comprehension ratings and tests matched that of the reading itself. Participants took approximately 40 minutes in the low-effort condition and approximately 50 minutes in the high-effort condition to complete the experiment.

RESULTS

First step of analysis, participants were categorized as low- or high-WMC readers on the basis of OST scores. Median split was used to assign participants. Participants who provided correct answers to less than 85% of the operations were considered to be
processing inappropriately (Conway et al., 2005). Thus, one participant was eliminated from the analyses. Results indicated that in the low-effort condition, 20 participants were categorized as high-WMC readers ($M = .43, SD = .13$), and 19 participants were categorized as low-WMC readers ($M = .19, SD = .08$). In the high-effort condition, 19 participants were categorized as high-WMC readers ($M = .43, SD = .11$), and 19 participants were categorized as low-WMC readers ($M = .19, SD = .06$).

**Comprehension Ratings and Test Performance.** We conducted a $2 \times 2$ ANOVA for comprehension ratings and performance on the comprehension test (see Table 1). The results of comprehension ratings showed no significant main effects and no interaction between high- and low-WMC readers ($F$s < 2.25, $p$s > .13). In contrast, the results of test performance showed no significant main effects of effort and no interaction between high- and low-WMC readers ($F$s < 1, $p$s > .32), but did show significant main effect of WMC [$F(1, 73) = 6.47, MSE = 2.22, p < .05$, partial $\eta^2 = .08$].

**Monitoring Accuracy.** Monitoring accuracy was operationalized as the Goodman-Kruskal correlation between comprehension ratings, which ranged from 1 to 7, and test performance, which ranged from 0 to 2. *Gamma* can vary from $-1.0$ to $+1.0$; $-1.0$ is perfect negative correlation, and $+1.0$ is a perfect positive correlation. *Gamma* is more appropriate as a measure of ordinal association than are other correlation coefficients (Gonzalez & Nelson, 1996; Nelson, 1984), and this measure has been used in many previous studies. We computed the *gamma* correlations for all participants, but could not compute the *gamma* for three participants who had given the same ratings or performances for all texts; another four participants whose *gamma* correlation was more than 2 standard deviations from the mean was considered an outlier. Thus, these participants were eliminated from analyses. As seen in Fig. 1, results of a $2 \times 2$ ANOVA showed no significant main effects ($F$s < 1) but did reveal a significant interaction between WMC and mental effort [$F(1, 66) = 7.65, MSE = .26, p < .01$, partial $\eta^2 = .10$]. Among low-WMC readers, the *gammas* were lower under the low-effort condition than under the high-effort condition [$F(1, 66) = 2.91, MSE = .26, p < .09$]. Among high-WMC

### Table 1. The Means for Comprehension Ratings and Test Performance.

<table>
<thead>
<tr>
<th>WMC</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Low-effort</td>
<td>Rating</td>
<td>5.12</td>
</tr>
<tr>
<td></td>
<td>Performance</td>
<td>5.85</td>
</tr>
<tr>
<td>High-effort</td>
<td>Rating</td>
<td>4.45</td>
</tr>
<tr>
<td></td>
<td>Performance</td>
<td>6.00</td>
</tr>
</tbody>
</table>

Note: Comprehension ratings range from 1 to 7 and comprehension test performances were scored as the total number of correct answers and ranged from 0 to 10.
readers, the *gammas* were higher under the low-effort condition than under the high-effort condition \([F(1, 66) = 4.79, \text{MSE} = .26, \; p < .05]\). Furthermore, under the low-effort condition, the *gammas* were higher for high-WMC readers than for low-WMC readers \([F(1, 66) = 4.87, \text{MSE} = .26, \; p < .05]\). Under the high-effort condition, the *gammas* were lower for high-WMC readers than for low-WMC readers \([F(1, 66) = 2.97, \text{MSE} = .26, \; p < .10]\).

We also conducted a *t*-test to evaluate whether *gamma* values were significantly different from zero; results for the low-effort condition showed that the *gamma* values for low-WMC readers did not differ from zero significantly \([t(18) = 1.73, \; p = .10]\), but results for other groups showed that the *gamma* values significantly higher than zero \((ts > 2, \; p < .05)\).

**DISCUSSION**

This study investigated how WMC and mental effort affected the accuracy with which text comprehension was monitored. The result of present study showed a significant interaction between WMC and effort. In the low-effort condition, high-WMC readers could monitor their comprehension more accurately than low-WMC readers did. In contrast, in the high-effort condition, low-WMC readers could monitor their comprehension more accurately than high-WMC did. Additionally, when effort of processing texts increased, monitoring of low-WMC readers improved, whereas that of high-WMC readers decreased. This result suggests that both WMC and mental effort affect monitoring accuracy, and thus, support our hypotheses.

However, in comprehension test performance, although the performance of high-
WMC readers was higher than that of low-WMC readers, we found no effect of effort. Thus, mental effort did not affect test performance. One possibility is that manipulation of text by deletion of letters has a minimal effect on comprehension in between-subject designs (Maki et al., 1990, Experiment 2). This may explain the absence of effect of effort in performance in the present study. However, reading deleted-letters texts, participants demanded to process complete fragments and understand texts simultaneously. In other words, participants were entailed dual task, which were the processing of complete fragments and texts, in the high-effort condition. Dual task, in general, demand more resources from participants (e.g., Nishizaki & Osaka, 2004). Given that the retrieval process is important for text comprehension (Nishizaki & Osaka, 2004), this retrieval process interfere in comprehension. Therefore, participants in high-effort condition might process more effortfully. Therefore, this finding of no effect of effort suggests that the allocation of resources for processing texts might not vary in each effort condition and that only the attention allocated to the meta-level differed.

We propose that the weight accorded cues might vary under conditions of increased mental effort. In cases of optimal effort, readers could allocate sufficient resources to the meta-level and also process cues concerning situation model and weight these cues accordingly. Because these cues affect performance on comprehension tests, they are efficient for inferring levels of comprehension (e.g., Anderson & Thiede, 2008; Thiede et al., 2003; Thiede et al., 2010). It is likely that this pattern of weighting corresponded to that used by high-WMC readers in the low-effort condition. In fact, their gamma correlations were higher than those typically reported, and were as high as that found in other studies after interventions in which readers generated key words or summaries following delays (Anderson & Thiede, 2008; Thiede & Anderson, 2003; Thiede et al., 2010). Furthermore, monitoring is most accurate under conditions of optimal effort (Weaver & Bryant, 1995). Thus, readers might weight cues concerning text representation under optimal effort condition.

In contrast, when readers need to allocate more resources to object-level because of increased effort to processing texts, they might be able allocate some resources to the meta-level and obtain some cues concerning text representation. However, such cues might be insufficient for inferring comprehension. Thus, readers would not be able to effectively weight cues according to their relevance to text representation, thereby reducing monitoring accuracy. If increased effort did not affect variations in the amount of resources allocated to the meta-level and the cues available for comprehension, the accuracy of high-WMC readers would not decrease.

However, when readers cannot allocate resources to the meta-level, they obtain fewer cues concerning text representation. As a result, effort itself, or surface elements, such as processing difficulty, is more available and subjectively salient cues. Given that individuals rely on salient cues (Koriat & Ma’ayan, 2005), they rely on these surface cues in high effort situation. But, if readers were to rely on surface cues alone, their accuracy could be limited. However, it is possible that surface-memory-based combination cues improve accuracy at some level (Thiede et al., 2010). If allocation of sufficient resources to the meta-level were all that mattered and only situation-model-based judgments
improved accuracy, the accuracy of low-WMC readers would not improve with increased effort delivering more resources to the object-level. Thus, perhaps readers rely on a combination of surface and memory-based cues, both of which influence how well they recall texts. As a result, the monitoring accuracy of low-WMC readers might have improved on this basis in the high-effort condition. This improvement, however, is not robust, because, compared with the cues concerning text representation, these cues are less effective. In fact, the improvement of accuracy for low-WMC readers and the difference of accuracy between low- and high-WMC readers under high effort condition were marginal.

In this way, variations in the weights accorded to cues during monitoring might lead to differences in monitoring accuracy, and this shift in weighting might be similar with regard to process but discrepant with regard to timing in high- and low-WMC readers. This discrepancy might reflect differences in attentional control ability. However, this possibility was not investigated directly by the present study, and it requires further investigation.

There are some limitations to the current research. First, there is an uncertainty regarding to the effectiveness of reading effort manipulation. In the experiment, although we manipulated processing effort, there was no difference in the test performance between each group. Additionally, reading time was not measured. Thus, it is unclear how this manipulation affected effort. Secondly, we used two comprehension test items for each text, because the texts were short. In contrast to previous studies, there were less numbers of items included in the experiment. Given that the number of test items affects gamma correlation (Weaver, 1990), it is possible that the smaller number of test items affected the gamma values. Although the results of the present study showed a robust interaction, these issues are important points, and thus, future research should be conducted with these issues in mind.

In conclusion, this study demonstrated that monitoring accuracy is affected by both WMC and mental effort. This result suggested that it is not the necessarily the case that the high-WMC readers predominate in monitoring of their comprehension. Furthermore, present study resolved inconsistent accounts which is dual processing assumption and increased processing hypothesis. We propose that the weighting of cues varies according to the quantity of cues pertaining to text representation. Future studies should examine this account and elaborate on the use of a metacognitive model for text comprehension.

REFERENCES


Rawson, K. A., Dunlosky, J., & Thiede, K. W. 2000. The rereading effect: Metacomprehension accuracy...


(Manuscript received 28 January, 2011; Revision accepted 4 June, 2012)