<table>
<thead>
<tr>
<th>Title</th>
<th>The role of executive function in integration process of arithmetic word problem solving: Focusing on the updating function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>森 賢隆</td>
</tr>
<tr>
<td>Editor(s)</td>
<td></td>
</tr>
<tr>
<td>Citation</td>
<td></td>
</tr>
<tr>
<td>Issue Date</td>
<td>2018</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://repository.osakafu-u.ac.jp/dspace/">http://repository.osakafu-u.ac.jp/dspace/</a></td>
</tr>
<tr>
<td>Rights</td>
<td></td>
</tr>
</tbody>
</table>
The Role of Executive Function in Integration Process of Arithmetic Word Problem Solving: Focusing on the Updating Function

Kanetaka MORI

Graduate school of humanities and social sciences

2018
Acknowledgements

This dissertation bases on two articles published.


- The Introduction of Chapter 1 and General discussions of Chapter 4 also bases on these two publications.

I would like to express my sincere gratitude to my supervisor Prof. Masahiko Okamoto for the continuous support of my Ph.D. study, for his patience, motivation, and immense knowledge.

Besides my supervisor, I would like to thank Prof. Wataru Ide and Prof. Shogo Makioka for their reviews, insightful comments, and encouragements.

I thank my school fellows for the stimulating discussions and for all the fun we have had in the study life.

I also would like to thank all participants for sparing some precious time for my study.

Finally, I express my heartfelt gratitude to my family, especially parents for watching over me.

Kanetaka Mori
2018
## Contents

1 Introduction 1
   1.1 Review of working memory ................................. 1
   1.2 Arithmetic problem solving and working memory .......... 4
      1.2.1 Study of arithmetic word problem solving ............ 4
      1.2.2 Cognitive functions required in arithmetic word problem solving 6
      1.2.3 Issues of previous studies and the purpose of present study . 7

2 The role of updating function in integration process of arithmetic word problem solving 9
   2.1 Importance of updating function in integration process of arithmetic word problem solving (Experiment 1) ........ 9
      2.1.1 Purpose ........................................... 9
      2.1.2 Methods ......................................... 10
      2.1.3 Results and discussions .............................. 12
   2.2 The role of updating function in constructing problem model (Experiment 2) .............................................. 17
      2.2.1 Purpose ........................................... 17
      2.2.2 Methods ......................................... 17
      2.2.3 Results and Discussion .............................. 19
   2.3 General discussions about the role of updating functions in integration process .............................................. 24
      2.3.1 Constructing a Problem Model and the Role of the Updating Function .............................................. 24
      2.3.2 Updating and Modality ................................. 25

3 Relationship between problem format and modality of updating functions 27
   3.1 Introduction .............................................. 27
      3.1.1 Factors relating to the dominance of each working memory domain .............................................. 28
      3.1.2 Purpose and hypothesis in this chapter .................. 28
   3.2 Area word problems and updating functions on phonological and visual domain (Experiment 3) ..................... 30
      3.2.1 Purpose .............................................. 30
      3.2.2 Methods ......................................... 30
      3.2.3 Results and discussions .............................. 33
3.3 Area problems in a figure format and updating functions on phonological and visual domain (Experiment 4) ........................................ 40
  3.3.1 Purpose ................................................................. 40
  3.3.2 Methods ............................................................... 40
  3.3.3 Results and Discussion ........................................... 41
3.4 General discussions about the modality of updating function in integration of problem solving ......................................................... 45

4 General discussion ...................................................... 47
  4.1 The updating function and integration process in arithmetic word problem solving ................................................................. 47
  4.2 Educational implications ............................................. 51

References ................................................................. 53

A Examples of arithmetic word problem in Experiment 1. ......................... 59
B Examples of arithmetic word problem in Experiment 2. ......................... 61
C Examples of arithmetic word problem in Experiment 3. ......................... 63
Chapter 1

Introduction

1.1 Review of working memory

When we perform some cognitive tasks, we should hold certain information in our memory (e.g., task relevant information, task-goal, etc.). Working memory plays an important role in retaining such information. According to Baddeley’s (1986) three-component model, working memory involves two sub-systems: a phonological loop, which stores verbal materials, and a visuospatial sketchpad, which houses visual information. These two components are coordinated by the central executive, a supervisory system concerned with attention. A revised version of the working memory model (Baddeley, 2000; Baddeley, Allen, & Hitch, 2011) incorporates an additional storage system called the episodic buffer. The buffer is added in order to account for the interaction of the verbal or visual domain with long-term memory. It is assumed to hold integrated episodes or chunks in a multidimensional code. Figure 1.1 shows the multicomponent working memory model (Baddeley, 2000).

The phonological loop is assumed to comprise a phonological store and an articulatory rehearsal mechanism. The phonological store holds information as phonological code. As phonologically similar items cause errors of memory, this phonological similarity effect is evidence of the existence of the phonological store. The information in the phonological store is maintained via the rehearsal mechanism, which is reflected in the fact that longer words reduce the performance of immediate recall because those require a greater rehearsal time. Although the phonological loop stores phonological codes, the loop stores not only auditory presented items but also visually presented ones. When material is presented visually, articulatory suppression extinguishes the phonological similarity effect and word length effect (Murray, 1968; Levy, 1971). Thus, visually presented items are translated into appropriate phonological codes through subvocal rehearsal. This loop is thought to play an important processing role in text comprehension. For example, the capacity of phonological short-term storage, which is measured by nonword repetition, is a good predictor of children’s reading problems (Gathercole & Pickering, 2000).

The visuospatial sketchpad is assumed to be a storage system capable of integrating visual and spatial information, which are acquired from vision, touch, language, or long-term memory, into a unitary visuospatial representation (Baddeley, 2007).
Logie (1995) proposed a visual cache as a passive storage system and an inner scribe as an active maintenance mechanism. These systems correspond to the phonological store and articulatory rehearsal mechanism in the phonological loop. This storage system is thought to play an important role in manipulating visual and spatial information. Pearson, Logie, and Gilhooly (1999) demonstrated that spatial tapping, which interferes with the visuospatial sketchpad, reduced successful mental synthesis. The recall of spatial mental representations derived from spatial descriptions relates to the performance of a Corsi span task in the backward version (Picucci, Gysselinck, Piolino, Nicolas, & Bosco, 2013; Meneghetti, Ronconi, Pazzaglia, & De Beni, 2014). The task is presented as an array of nine blocks, and the experimenter points to them in a predetermined order. The participant’s task is to point to the blocks either in the same order as the experimenter or in the reverse order. The series starts with only two blocks, and progresses to more blocks until the child fails to correctly copy the order on two consecutive trials with the same number of blocks. The Corsi span task is assumed to engage the visuospatial sketchpad.

The central executive is an attentional control system with limited capacity. The central executive system is thought to be responsible for managing the flow of information in and out of the various stores (Cowan, 2017). In working memory assessment, especially in the educational context, four aspects of working memory are assessed: verbal short-term memory, visual short-term memory, verbal working memory, and visual working memory (Alloway, Gathercole, & Pickering, 2006; Alloway, Gathercole, Kirkwood, & Elliott, 2008). The short-term memory aspects are assessed by tasks that simply require information, such as a digit span task for the
1.1. Review of working memory

verbal domain or a block recall task for the visual domain, to be retained. In the
digit span task, some digits (e.g., 5, 4, 7) are presented. We must recall the digits in
the same sequence of presentation (e.g., 5, 4, 7). The presentation is often verbal. On
the other hand, the working memory aspects are assessed via tasks that require the
maintenance and active processing of information. For example, a backward digit
span task requires reversing the order of the holding information (e.g., 7, 4, 5). This
task is an index of verbal working memory. The verbal short-term memory reflects
the function of the phonological loop. This verbal working memory is assumed to be
a performance when we process information in the verbal short-term memory using
processing resources of the central executive function. Similarly, the visual work-
ing memory relates to the visuospatial sketchpad and the central executive function.
Thus, the abilities of working memory are tapped by tasks requiring the central ex-
ecutive system.

Working memory is thought to be a limited-capacity system that provides the
temporary storage and manipulation of information necessary for performing a wide
range of cognitive activities (Baddeley, 2012). Daneman and Carpenter (1980) per-
formed one of the earliest studies to indicate the importance of the active aspect of
memory (i.e., working memory). They investigated whether reading comprehension
measures correlate with certain memory span tasks (e.g., digit span and reading
span). The reading span is indexed as the number of final words recalled when par-
ticipants remember the last word of sentences when deciding the of validity of the
sentences. The span is assumed to tap the working memory capacity (WMC) rather
than the capacity of short-term memory because this task requires participants to
maintain and process some pieces of information simultaneously. The results indi-
cate that performance on the reading span test was correlated with reading ability,
while the digit span, which is a measure of capacity of short-term memory, did not
show such a correlation. The WMC has been reported to show relationships with
a broad range of cognitive activities: reading comprehension (Daneman & Merikle,
1996) and reasoning (Kane & Engle, 2002), for example.

Measures that tap working memory require active processing. This implies a re-
lationship with some roles of the central executive function. However, functions of
the central executive have been left unspecified for a long time. Recently, some re-
searchers have suggested a range of functions of the central executive and attempted
to analyze the central executive into subcomponents (Baddeley, 1996; Miyake et al.,
2000). Miyake et al. (2000) provide evidence that the executive functions include
at least three functions, described as inhibiting, shifting, and updating. Inhibiting
is the function of inhibiting dominant, automatic, and prepotent responses. This
function was tapped by a stroop task, an antisaccade task, and a stop-signal task.
Shifting is the function of switching back and forth flexibly between tasks or mental
sets. A color-shape task, a number-letter task, and a category-switch task have been
related to this function. Finally, updating is the function of monitoring incoming
information for relevance to the task at hand and then appropriately updating old,
no longer relevant information by replacing it with new, more relevant information. This function was tapped by a letter memory task, a keep-track task, and a spatial n-back task. The relationship between executive functions and complex executive tasks was also investigated (Miyake et al., 2000). They found that each complex executive task relates differently to these three executive functions. The Wisconsin card sorting task and Tower of Hanoi were found to be related to the shifting and inhibiting functions, respectively; random number generation related to the inhibition and updating functions; and operation span, which includes maintenance and processing information, is related to the updating function. Therefore, each cognitive function supports a respective range of our daily complex cognitive tasks. For example, when a phone rings as we are driving a car, we should continue to drive by inhibiting the action of responding to the incoming call; if the updating function malfunctions, we could sprinkle salt repeatedly while cooking even if it is no longer necessary; and we might need recourse to the shifting function when changing conversation partners from colleagues to customers.

As we described above, the working memory and particularly executive functions play a number of important roles in our cognitive activities. It is important to determine which central executive functions relate to a given cognitive activity, as such investigations could clarify cognitive activities. Also, such an investigation could provide evidence to foster our understanding of the theoretical model.

1.2 Arithmetic problem solving and working memory

1.2.1 Study of arithmetic word problem solving

An arithmetic word problem is defined as an arithmetic exercise presented in text rather than mathematical notation. Problem solvers have to solve the word problem using mathematical concepts and procedures. They read the sentences of the problem and comprehend it, then provide the numerical expressions to solve the problem by themselves. These cognitive activities are also supported by working memory. Recently, many studies have investigated how working memory relates to arithmetic word problem solving. To solve an arithmetic word problem, problem solvers must carry out multiple processes: they must read the problem, store the relevant information, plan out the solution, and perform calculations according to the plan. Working memory is used as a mental workspace for all these processes in word problem solving.

Before we discuss the relationship between working memory and arithmetic word problem solving, we review some studies of arithmetic word problem solving. Many studies have attempted to understand children’s struggles in solving arithmetic word problems. Some studies of arithmetic word problems have investigated whether children understand the language of word problems. Because word problems are presented linguistically, children fail to solve them if they fail to read
1.2. Arithmetic problem solving and working memory

<table>
<thead>
<tr>
<th>Necessary-information problem</th>
<th>Extraneous-information problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>There are 5 pencils.</td>
<td>There are 23 apples.</td>
</tr>
<tr>
<td>There are 1.6 times more pens than pencils.</td>
<td>There are 3 times more oranges than apples.</td>
</tr>
<tr>
<td>How many pens are there?</td>
<td>There are 8 times more grapes than apples.</td>
</tr>
<tr>
<td>How many oranges are there?</td>
<td>How many oranges are there?</td>
</tr>
</tbody>
</table>

Note. The underlined sentence is extraneous information.

them. Indeed, Lewis and Mayer (1987) reported that children often provided incorrect answers based on correct calculations performed from incorrect understanding. Furthermore, Schumacher and Fuchs (2012) demonstrated that understanding the relational terminology in a word problem is related to improved performance in solving compare problems. This finding suggests that understanding the terminology and language used in the word problem is needed to solve the problem correctly. Thus, some failures in word problem solving have been attributed to the children’s language skills.

However, children sometimes produce erroneous solutions even in cases where they understood each statement correctly. Muth (1984, 1992) assigned a problem-solving task in which the problem included extraneous information, that is, information not required to solve the problem. We refer to such a problem as an extraneous-information problem and to a problem that includes only necessary information as a necessary-information problem. Table 1.1 shows examples of these problems. Her experiment showed that a child who could solve a necessary-information problem was likely to fail to solve an extraneous-information problem. Although the extraneous information did not change the syntactic complexity or the expression of a problem, it increased the errors produced in problem solving.

According to studies of the cognitive process of arithmetic word problem solving (Mayer & Hegarty, 1996; Polya, 1957; Kintsch, 1992), the problem-solving process can be divided into a comprehension phase and a problem-solving phase. The comprehension phase can also be divided into a translation phase and an integration phase, and the problem-solving phase can be divided into a planning phase and an execution phase (Figure 1.2). In the translation phase, problem solvers need to translate each statement of a problem into a mental representation, and must integrate each representation into a problem model in the integration phase. Problem solvers use numerical expressions to solve the problem based on the problem model employed. The process that results in the formation of a particular model is important because the solution strategy depends on it. Cummins, Kintsch, Reusser, and Weimer (1988) report that most errors in word problems are due to inadequate integration of the problem in the integration phase. The findings of Muth (1984, 1992) also indicate the importance of the integration process because extraneous information increases the difficulties of the integration process but does not interfere with the translation process. This finding suggests that whether a problem-solving
Chapter 1. Introduction

FIGURE 1.2: Cognitive processes in solving arithmetic word problems (based on Mayer & Hegarty, 1996)

An attempt is successful depends on the quality of the solver’s problem model based on integration process. In solving arithmetic word problems, the most difficult and important phase seems to be the integration process.

1.2.2 Cognitive functions required in arithmetic word problem solving

A large body of research has shown that working memory is an important cognitive component in solving arithmetic word problems. Several studies have revealed the importance of the central executive system in arithmetic word problem solving (Andersson, 2007; Fuchs et al., 2010; Lee, Ng, Ng, & Lim, 2004; Meyer, Salimpoor, Wu, Geary, & Menon, 2010; Swanson, 2004, 2006; Swanson & Sachse-Lee, 2001). For example, Swanson (2004) reports that the central executive contributes to accuracy in arithmetic word problem solving when phonological processing, fluid intelligence, and reading comprehension are controlled for. Furthermore, Lee et al. (2004) demonstrate that both the phonological loop and the visuospatial sketchpad contribute to performance on arithmetic word problems via literacy and performance IQ, respectively. Meanwhile, the central executive contributes to performance on this task both directly and indirectly. These studies thus indicate that the central executive system plays an important role in word problem solving. That is, in arithmetic word problem solving, it is important to perform some active processing of information, such as selecting and relating information or providing a numerical expression, via an attentional control system with limited capacity.

Previous studies, however, had not provided clear evidence on the specific contribution of the central executive to arithmetic word problem solving because the central executive seems to have multiple functions (Baddeley, 1996; Miyake et al., 2000). As described above, Miyake et al. (2000) provide evidence that the functions of the central executive include at least three executive functions, described as inhibition, switching, and updating. Recently, studies have investigated the relationship between these three functions and arithmetic word problem solving. Agostino, Johnson, and Pascual-Leone (2010) report that updating proficiency is the best predictor of accuracy in arithmetic word problem solving, based on structural equation
modeling. Passolunghi and Pazzaglia (2005) also provide evidence that the updating function is involved in the process of solving arithmetic word problems. They demonstrate that children who exhibit better mathematical performance also exhibit a higher updating performance. These studies suggest that the updating function is the most important of the three executive functions in solving an arithmetic word problem.

1.2.3 Issues of previous studies and the purpose of present study

Studies of arithmetic word problem solving have suggested the importance of the updating function and the integration process. However, the relationship between the updating function and the integration process has not been clearly investigated. For example, Muth (1984, 1992) revealed that the integration process is important and that extraneous information increases the difficulty of the process, but did not indicate how this effect occurs or which cognitive functions related to a successful integration process. On the other hand, investigations of the relationship between working memory and arithmetic word problem performance reveals that the updating function is important in solving arithmetic word problems. However, the role of the updating function in solving arithmetic word problems also remains unclear. An important assumption generated from previous research into the relationship between working memory and word problem solving is that the updating function may play a critical role in the integration process. In the integration process, one needs to continually integrate incoming information to form a problem model (Kintsch, 1992). When new information enters working memory, it changes the model and the old problem model should be replaced by a newer one. This means that problem solvers might have to update the nature of their problem model sequentially when they are carrying out a word problem-solving task. A failure in updating results in an inappropriate problem model and an incorrect answer. Therefore, we hypothesize that updating during integration is crucial for word problem solving.

Evidence to support this prediction comes from Kotsopoulos and Lee (2012), who examined the phase and nature of errors occurring in word problem solving. They video-recorded students talking aloud while completing homework. Their speech coding identified which of the four problem-solving phases appeared problematic for students and which single executive function played the most important role in explaining their difficulties. Challenges in updating occurred when students had difficulties evaluating information and appropriately editing it in the light of more relevant information in ways that would allow them to proceed to the next problem-solving phase. This research revealed that most of the errors occurred in the integration process, and were caused by a failure of the updating function in integration. Integration might be a main process in solving word problems, and updating is a key cognitive function in this process. However, these results were obtained from subjects’ verbal reports; as a result, the underlying cognitive processes were not clearly
Chapter 1. Introduction

reflected. There is no direct evidence that their errors were due to the updating function, because Kotsopoulos and Lee (2012) did not objectively measure students’ updating function. Thus, the relationship between the updating function and the integration process needs to be further investigated empirically.

This dissertation seeks to comprehend how the updating function relates integration process of solving arithmetic word problems. First, in Chapter 2, the relationship between the updating function and the integration process was investigated. Experiment 1 revealed the importance of the updating function in the integration process. Experiment 2 was conducted to examine the role of the updating function to constructing the problem model in integration. In Chapter 3, the dependence of the relationship between the updating modality and the integration process on problem format or category of problem was investigated. Experiment 3 showed the relationship when solving area word problems, and the relationship when solving area problems in a figure format was investigated in Experiment 4.
Chapter 2

The role of updating function in integration process of arithmetic word problem solving

2.1 Importance of updating function in integration process of arithmetic word problem solving (Experiment 1)

2.1.1 Purpose

The purpose of this experiment was to reveal the importance of the updating function in integration when solving an arithmetic word problem. Investigating the relationship between the updating function and the integration process should provide insights into how word problems are solved and what process we must be careful to teach children who have lower facility in updating.

The contribution of working memory to arithmetic word problems is not stable in elementary school students (Rasmussen & Bisanz, 2005; Meyer et al., 2010). Moreover, if problem solvers do not acquire appropriate language skills, they cannot solve arithmetic word problems, and in that case it is of course not possible to examine the contribution of working memory to arithmetic word problem solving. Because working memory and language skills are confounded in children, we used undergraduate students with stable working memory, sufficient language skills, and related problem-solving schema.

Hegarty, Mayer, and Green (1992) and Okamoto (1999) examined the nature of word problems that are challenging for children and adults by examining the reading times required for integration. They found that both children and adults require more time for integration when solving extraneous-information problems. Thus, the extraneous information could increase the difficulty of integration (Muth, 1984, 1992). These results suggest that the use of extraneous-information problems when measuring reading times are useful for examining the integration phase in solving word problems. To explore the relationship between the updating function and the integration process, we measured reading times when undergraduate students were solving necessary- and extraneous-information problems.
According to Kintsch (1992), it is necessary to elaborate a problem model while solving a problem. This process of making a problem model involves updating an older problem model to a newer problem model in working memory. The process might depend more strongly on the updating function when solving an extraneous-information problem than when solving a necessary-information problem. There is a possibility that the updating function supports the reduction of the increased difficulty of integration due to extraneous information. Therefore, we hypothesized if the updating function aids an efficient integration process, a problem solver whose updating function has a higher level of facility might be less influenced by the increased difficulty of integration.

Our predictions were as follows: (1) An extraneous-information problem requires more integration time than a necessary-information problem; and (2) the effect of extraneous information on integration time is smaller in higher-level-updating problem solvers.

2.1.2 Methods

Participants

Participants were 78 undergraduate students (42 males and 36 females) in Japan with an average age of 19.54 years. Most participants were recruited from a lecture of Introduction to Psychology in 2012 and 2015. They participated voluntarily, but were compensated with course credits or a book coupon for 500 yen. Some participants were recruited personally from acquaintances who were not very familiar to the researcher. All participants were native Japanese speakers and were enrolled in the following courses: Social Science, Human Science, Engineering, Science, and Health Science. Math-related courses were Engineering and Science. Participants in other courses also had sufficient mathematical achievement, as demonstrated by their passing of the National Center Test for University Admissions in Japan. Sample problems of this test were reported in Wu (1993). Mathematics in this test includes vectors and functions.

Apparatus

Stimulus presentation and response recording were controlled by Matlab with the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) on an Apple iMac 21.5-inch display and an original USB response key box with three buttons.

Tasks and Procedures

General Procedures. Participants were asked to perform an arithmetic word problem-solving task and then two updating tasks. The order of the phonological and visual updating tasks was counterbalanced across participants. Participants
received instructions, which were displayed on the monitor, prior to undertaking each task, and they could ask the experimenter any question. The total experimental time was approximately forty minutes. The procedures used in Experiment 1 were approved by the Research Ethics Committee of the Graduate School of Humanities and Social Sciences at Osaka Prefecture University.

**Arithmetic word problem-solving task.** Participants were asked to read a word problem and select a formula to solve it. Sentences were displayed one by one. Participants could manipulate the displayed sentence, moving backward or forward by pressing the left or right key on the response box. They were required to press the center key once they had comprehended the problem. After pressing this key, three expressions appeared, and participants selected the correct expression. To analyze the comprehension process in arithmetic word problem solving, we calculated three types of reading times (Figure 2.1). The time participants took to read a word problem sentence until expressions were displayed was the “whole reading” time. Time taken from onset of a problem to the end of the question statement’s presentation was “translation” time. Time taken from the question statement’s initial presentation to the expressions’ presentation was “integration” time. This reflected the integration process in solving a word problem. Hegarty et al. (1992) and Okamoto (1999) measured the time from a question statement’s presentation to selecting a numerical expression to solve a problem. This reflects not only the integration process but also the planning process. However, we must distinguish integration time from planning time. Thus, in this study, integration time was defined as the time from the onset of question presentation to the end of reading a problem. Planning time was defined as the reaction time needed to select the correct expression. We prepared two types of arithmetic word problems: necessary-information and extraneous-information problems. Table 1.1 shows an example of each. A necessary-information problem consisted of three sentences—all needed to form an expression. The extraneous-information problem contained four sentences—one of which was unnecessary for forming an expression. We used 16 arithmetic word problems. Half of the problems were necessary-information problems and the others were extraneous-information problems. Four problems were used for practice trials.

**Updating tasks.** We used two updating tasks: (1) a letter memory task (adapted from Miyake et al., 2000) for the phonological domain, and (2) a visual n-back task (adapted from Agostino et al., 2010) for the visual domain. The letter memory task measured phonological updating in working memory. In this task, letters were presented serially on a computer screen at a rate of 2000 ms per letter. Participants were required to rehearse aloud the last four letters throughout by dropping the fifth letter and adding the most recent one. We recorded participants’ verbal recall. The numbers of letters in the stimulus sets were 5, 7, 9, and 11. Each set was presented four times, and each participant completed 16 trials. If participants could accurately
name aloud all letters in each set, the trial was classified as correct. The phonological updating score was the percentage of correct trials. The visual n-back task measured visual updating of working memory. Three dots were presented on a computer screen. Participants were required to decide whether the current pattern was the same as the pattern n trials before. Their decisions were recorded by pressing a key that changed the presentation pattern. In each n-back task, there were 54 test trials following 12 practice trials. We used 20 matched test trials and 34 mismatched trials. We used the 1-back and 2-back tasks. Each score was calculated by the following formula: (proportion of correct match + proportion of correct mismatch).

2.1.3 Results and discussions

One participant was excluded from the analysis because of low accuracy (68%) in selecting a correct expression. Thus, data from 77 participants were analyzed. Whole reading, translation, integration, and planning times were calculated after excluding the error trials. Trials were excluded if an individual’s whole reading time exceeded the mean ±2 standard deviations (SD) in the participant. For analysis of planning time, trials over the mean ±2 SD in a participant for planning time were excluded. All reading times were standardized by each subject because there were individual differences in times (whole reading, translation, and integration times). Table 2.1 shows basic statistics for all measures. The phonological updating score was measured by a letter memory task, and the visual updating score was obtained using a visual 2-back task.
2.1 Importance of updating function in integration process of arithmetic word problem solving (Experiment 1)

### Table 2.1: Basic statistics for whole reading, translation, integration, and planning times; accuracy in selecting an expression; and updating functions in Experiment 1.

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole reading time</td>
<td>1073.10</td>
<td>3959.68</td>
<td>3621.33</td>
<td>23049.50</td>
</tr>
<tr>
<td>Translation time</td>
<td>8949.21</td>
<td>3771.67</td>
<td>2588.80</td>
<td>23049.50</td>
</tr>
<tr>
<td>Integration time</td>
<td>3944.38</td>
<td>1629.80</td>
<td>1275.00</td>
<td>9757.83</td>
</tr>
<tr>
<td>Planning time</td>
<td>1781.65</td>
<td>573.97</td>
<td>804.00</td>
<td>4382.50</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.94</td>
<td>0.07</td>
<td>0.75</td>
<td>1.00</td>
</tr>
<tr>
<td>Phonological updating</td>
<td>0.62</td>
<td>0.25</td>
<td>0.06</td>
<td>1.00</td>
</tr>
<tr>
<td>Visual updating</td>
<td>0.84</td>
<td>0.1</td>
<td>0.46</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Accuracies for selecting a correct expression were analyzed using a paired sample t-test for word problem type (necessary-information problem vs. extraneous-information problem). This analysis showed a significant difference in accuracy between word problem types ($t(76) = 2.98, p < .01$), which indicated that participants made more errors with extraneous-information problems than with necessary-information problems. Similarly, whole reading time was analyzed using a paired sample t-test for word problem type. This analysis also showed a significant difference ($t(76) = 22.69, p < .001$), indicating that participants had more difficulty and needed more time to comprehend an extraneous-information problem than a necessary-information problem.

### Individual differences of updating in solving word problems

To estimate the contribution of the updating function in solving word problems, we used a multiple regression analysis, with translation, integration, and planning times as dependent variables and problem type, phonological updating score, visual updating score, and these interactions as independent variables. Before the analysis, dependent variables were standardized, and the phonological updating score and visual updating score were also standardized. Problem type was coded as a dummy variable, with 0 indicating a necessary-information problem and 1 an extraneous-information problem. In these analyses, phonological and visual updating scores of each participant were entered repeatedly for problem type. Table 2.2 shows results of regression analysis for each dependent variable.

Analysis of translation time showed that phonological updating did not show any significant contribution. This result suggested that the phonological updating function was not important in translation. Furthermore, word problem type and interaction of visual updating and word problem type were significant ($\beta = 1.75, t(146) = 21.92, p < .001; \beta = -0.18, t(146) = -2.20, p < .05$). These results indicated that although translation time was longer for extraneous-information problems than for necessary-information problems, higher visual updating reduced this tendency. The requirement for more translation time with extraneous-information
problems was mostly due to the extraneous-information sentence. Higher visual updating solvers might be faster at encoding sentences than lower visual updating solvers. This effect was observed especially in extraneous-information problems because these problems included an additional sentence to be encoded.
### Table 2.2: Standardized partial regression coefficients of regression analysis for translation, integration, and planning times in Experiment 1.

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Translation time</th>
<th>Integration time</th>
<th>Planning time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem type</td>
<td>1.75***</td>
<td>1.43***</td>
<td>0.03</td>
</tr>
<tr>
<td>Phonological updating</td>
<td>−0.06</td>
<td>0.09</td>
<td>−0.07</td>
</tr>
<tr>
<td>Visual updating</td>
<td>0.10</td>
<td>−0.08</td>
<td>0.07</td>
</tr>
<tr>
<td>Phonological updating × visual updating</td>
<td>0.04</td>
<td>0.02</td>
<td>0.05</td>
</tr>
<tr>
<td>Phonological updating × problem type</td>
<td>0.08</td>
<td>−0.23*</td>
<td>−0.12</td>
</tr>
<tr>
<td>Visual updating × problem type</td>
<td>−0.18*</td>
<td>0.16</td>
<td>0.05</td>
</tr>
<tr>
<td>Three-way interaction</td>
<td>−0.07</td>
<td>−0.04</td>
<td>−0.04</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.76</td>
<td>0.51</td>
<td>−0.02</td>
</tr>
<tr>
<td>$F(7, 146)$</td>
<td>69.75***</td>
<td>23.80***</td>
<td>0.58</td>
</tr>
</tbody>
</table>

**Note.** ***$p < .001$, **$p < .01$, *$p < .05$*
In integration reading time, regression analysis revealed a significant contribution to word problem type ($\beta = 1.43, t(146) = 12.63, p < .001$). This further revealed a significant contribution for the interaction of phonological updating and word problem type ($\beta = -0.23, t(146) = -2.01, p < .05$), but not for phonological updating. These results indicated that in necessary problems, the phonological updating function did not contribute to integration, while in extraneous problems, the higher phonological updating function reduced integration time. These results support our hypothesis that the updating function plays an important role in integration process.

Analysis for planning time did not show any significant contribution. Even problem type was not significant. This indicated that planning time was the same for necessary-information and extraneous-information problems, suggesting that problem solvers completed the integration process before selecting an expression in this experiment.

In summary, in their integration, lower phonological updating solvers were more strongly affected by extraneous information than higher phonological updating solvers. The integration time difference might be caused by differences in their updating functions. Specifically, findings regarding integration time indicate that the integration process depends on the updating function. If the integration process differs by the updating function, the problem model that arises from the integration process can be different. However, the nature of the problem model resulting after the integration process was not clear in this experiment and requires further investigation.

According to Kintsch (1992), the integration process includes the elaboration of the problem model. In this elaboration, one forms an appropriate problem model that activates only the information required to solve the problem. However, extraneous information is not activated in such a clear problem model. One possibility to account for the difference in phonological updating abilities is that the updating function can help elaborate the problem so as to construct a clear problem model.

An alternative explanation is that working memory capacity (WMC) is also available. The WMC is indexed by a task score of how many items are being held while processing information. For example, in a reading span task the participant is required to verify a sequence of sentences and then recall the last word of each. The WMC is indexed as the maximum number of words that the participant can successfully verify and recall. This means that problem solvers who have a higher WMC can retain more information when solving a problem. Problem solvers with high WMC could thus more easily solve arithmetic word problems than those with low WMC (e.g., Rasmussen & Bisanz, 2005; Geary, 2004). If this advantage of high-WMC problem solvers is a consequence of their being able to hold more information in memory, such problem solvers form a model that includes all the information in the problem statement, and the problem model might use the extraneous information to ensure high activation. Investigation into the specific nature of activation based
2.2. The role of updating function in constructing problem model (Experiment 2)

2.2.1 Purpose

The purpose of the experiment was to reveal the nature of the problem model as a function of updating. We used a lexical decision task for this purpose. Participants were required to decide whether a presented stimulus was a word or not. This type of decision is made more rapidly if the same word has been previously presented, because the stimulus word’s representation is already activated. A lexical decision task is useful for investigating the representation’s activation in working memory.

We predicted response times (RTs) in the lexical decision task. If problem solvers with high updating formed a clear problem model, RTs for necessary-information words could be faster than those for extraneous-information words. In lower updating solvers, representations of both extraneous- and necessary-information words could still be activated. This is because their problem model could not be updated and they were likely to form a problem model including extraneous information. Therefore, in comparison to a novel word, the lexical decision for these two words could be more rapid. The RTs for extraneous- and necessary-information words might not differ for lower updating solvers, while RTs for necessary-information words could be faster than those for extraneous-information words in higher updating solvers. Investigating RTs for these three types of words in a lexical decision task would enable us to explore the nature of the problem model.

2.2.2 Methods

Participants

In Experiment 2, 73 undergraduate and graduate students in Japan (29 males and 44 females) participated. Their average age was 19.59 years. Most participants were recruited from a lecture of Introduction to Psychology in 2013 and 2015. Although they participated voluntarily, they received course credits or a book coupon worth 500 yen. Some participants were recruited personally from the researcher’s acquaintances who were not very familiar. All participants were native Japanese speakers and were enrolled in the following courses: Social Science, Human Science, Engineering, Science, and Health Science. Math-related courses were Engineering and Science. None had participated in Experiment 1.

Tasks and Procedures
General Procedures. As procedures in Experiment 1, participants were asked to perform the arithmetic word problem task and then two updating tasks. The order of the phonological and visual updating tasks was counterbalanced across participants. The procedures used in Experiment 2 were approved by the Research Ethics Committee of the Graduate School of Humanities and Social Sciences at Osaka Prefecture University.

Arithmetic word problem-solving task. To explore activation immediately after the integration process, we assigned the lexical decision task between reading a problem and selecting a correct expression.

First, participants were instructed to read a problem until they comprehended it. In the first phase, each sentence of a problem was displayed individually so that participants were free to choose which sentence they wanted to read. After participants pressed a key that indicated they had understood the problem, they were given the lexical decision phase. In this phase, two stimuli (consisting of a word and a nonword) were presented horizontally on a computer screen. Participants were required to select a word from these two stimuli by pressing a key (the left key for the stimulus on the left and the center key for the stimulus on the right). After the lexical decision phase, two expressions were presented horizontally on the computer screen. One of the two expressions was the correct solution to the arithmetic word problem. Participants were required to select the correct expression by pressing a key.

A set of 24 arithmetic word problems was developed. Half of the problems used extraneous information which was selected in a pseudo-random manner. This selection was counterbalanced across participants. Extraneous-information problems had an assignment, two related sentences, and a question. The two sentences included an extraneous sentence, which was unnecessary for solving the problem, and a relevant sentence, which was necessary for solving it. A variable noun in each sentence was called the extraneous-information word or the necessary-information word for this study’s purposes. Although a typical lexical decision task requires a decision about a word or a nonword in one presented stimulus, the lexical decision task in this study required participants to select a word stimulus from two letter strings. After the lexical decision task, selection of a correct expression from two expressions was immediately presented. We used this lexical decision task to reduce cognitive load from task switching. In the lexical decision task, the presented words had three conditions: (1) the necessary-information word condition, (2) the extraneous-information word condition, and (3) the novel word condition. Based on word frequency norms for Japanese (Amano & Kondo, 1999), 72 words written in Hiragana were collected. Half of the words had two characters and the other half had three. We prepared 24 sets of three words that corresponded to three conditions. A set of three words had the same length of characters and similar frequency. By linking two-letter syllables not normally associated with each other (Umemoto,
2.2. The role of updating function in constructing problem model (Experiment 2)

### Table 2.3: Basic statistics for reading time in each phase, accuracy for selecting an expression, reaction time, accuracy for lexical decision, and updating functions in Experiment 2.

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole reading time</td>
<td>10054.06</td>
<td>3311.47</td>
<td>3511.25</td>
<td>21838.50</td>
</tr>
<tr>
<td>Translation time</td>
<td>8633.82</td>
<td>3114.11</td>
<td>1994.92</td>
<td>15727.18</td>
</tr>
<tr>
<td>Integration time</td>
<td>3220.31</td>
<td>1581.40</td>
<td>941.75</td>
<td>11158.83</td>
</tr>
<tr>
<td>Planning time</td>
<td>1117.98</td>
<td>345.73</td>
<td>491.00</td>
<td>2620.75</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.98</td>
<td>0.03</td>
<td>0.83</td>
<td>1.00</td>
</tr>
<tr>
<td>RTs for lexical decision</td>
<td>863.71</td>
<td>214.14</td>
<td>513.75</td>
<td>1929.25</td>
</tr>
<tr>
<td>Accuracy for lexical decision</td>
<td>1.00</td>
<td>0.01</td>
<td>0.92</td>
<td>1.00</td>
</tr>
<tr>
<td>Phonological updating</td>
<td>0.65</td>
<td>0.22</td>
<td>0.19</td>
<td>1.00</td>
</tr>
<tr>
<td>Visual updating</td>
<td>0.92</td>
<td>0.11</td>
<td>0.49</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Morikawa, & Ibuki, 1955), 24 nonwords were created. Nonwords comprising three letters were formed by adding one letter to two-letter syllables. Furthermore, six words and four nonwords were prepared for four practice problems.

**Updating tasks.** A letter memory task and a visual n-back task were used to index the updating function. The visual n-back task was devised based on two main features from Experiment 1 (adapted from Friedman et al., 2008). The first feature was the presentation duration. In Experiment 1, the stimulus was displayed until participants pressed a key. To ensure more updating in Experiment 2, the stimulus was displayed for 1000 ms, followed by a blank screen for 1500 ms. Participants pressed a key indicating whether or not the stimulus was same as the one that had been displayed as the n-previous stimulus in that blank. The second feature was the presentation stimulus. One black square and nine white squares were presented on a computer screen. As in Experiment 1, the phonological updating score was the percentage of correct responses and the visual updating score was the 2-back task score.

2.2.3 Results and Discussion

**Analysis of word problem-solving process in problems containing extraneous information.**

Whole reading, translation, integration, and planning times were analyzed after excluding error trials to select an expression. Trials were also excluded for analysis if whole reading time was over the mean $\pm 2 SD$ for each participant. For analysis of planning time, trials over the mean $\pm 2 SD$ in a participant for planning time were excluded. One participant was excluded because she had many errors in the lexical decision task and struggled to select the correct expression (correct responses were 92% in the lexical decision task and 92% in selecting an expression). Table 2.3 shows the basic statistics for all measures.
Accuracy in selecting correct expressions was analyzed using a paired sample t-test for word problem type (necessary vs. extraneous). This analysis showed a significant difference between word problem types ($t(71) = 2.65, p < .01$). Furthermore, whole reading time was analyzed using a paired sample t-test for word problem type. This analysis showed that the difference between whole reading times for word problem types was significant ($t(71) = -25.03, p < .001$). These results indicated that extraneous-information problems are more difficult than necessary-information problems for participants’ comprehension and solution.

Same as that in Experiment 1, translation, integration, and planning times were analyzed with the multiple regression method. Table 2.4 shows results of regression analysis. Word problem type, phonological updating, visual updating, and these interactions were used as independent variables. The results showed a significant contribution of word problem type to translation time ($\beta = 1.73, t(136) = 20.57, p < .001$). For integration reading time, the results revealed that the contribution of word problem type ($\beta = 1.66, t(136) = 17.82, p < .001$) and interaction of phonological updating and word problem type ($\beta = -0.20, t(136) = -2.17, p < .05$) were significant. These results indicated that lower phonological updating solvers were more influenced by extraneous information than higher phonological updating solvers.

Analysis of planning time showed that the contribution of phonological updating was significant ($\beta = -0.40, t(136) = -3.58, p < .001$). The results indicated that problem solvers who had the higher phonological updating were likely to select a correct expression faster. The results of the word problem-solving process in Experiment 2 were consistent with those in Experiment 1.
2.2. The role of updating function in constructing problem model (Experiment 2)

TABLE 2.4: Standardized partial regression coefficients of regression analysis for translation, integration, and planning times in Experiment 2.

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Translation time</th>
<th>Integration time</th>
<th>Planning time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word problem type</td>
<td>1.73***</td>
<td>1.66***</td>
<td>0.26</td>
</tr>
<tr>
<td>Phonological updating</td>
<td>-0.02</td>
<td>0.10</td>
<td>-0.40***</td>
</tr>
<tr>
<td>Visual updating</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.10</td>
</tr>
<tr>
<td>Phonological updating × visual updating</td>
<td>0.02</td>
<td>0.04</td>
<td>-0.02</td>
</tr>
<tr>
<td>Phonological updating × word problem type</td>
<td>0.04</td>
<td>-0.20*</td>
<td>-0.01</td>
</tr>
<tr>
<td>Visual updating × word problem type</td>
<td>0.09</td>
<td>-0.02</td>
<td>-0.04</td>
</tr>
<tr>
<td>Three-way interaction</td>
<td>-0.04</td>
<td>-0.08</td>
<td>-0.04</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.75</td>
<td>0.69</td>
<td>0.14</td>
</tr>
<tr>
<td>$F(7, 136)$</td>
<td>60.92***</td>
<td>46.07***</td>
<td>4.27***</td>
</tr>
</tbody>
</table>

Note. *** $p < .001$, ** $p < .01$, * $p < .05$
Table 2.5: Standardized partial regression coefficients of regression analysis for RTs in lexical decision tasks in Experiment 2

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Necessary condition</td>
<td>−1.34***</td>
</tr>
<tr>
<td>Extraneous condition</td>
<td>−0.91***</td>
</tr>
<tr>
<td>Phonological updating</td>
<td>−0.45***</td>
</tr>
<tr>
<td>Visual updating</td>
<td>0.25*</td>
</tr>
<tr>
<td>Phonological updating × visual updating</td>
<td>0.01</td>
</tr>
<tr>
<td>Phonological updating × necessary condition</td>
<td>0.22</td>
</tr>
<tr>
<td>Phonological updating × extraneous condition</td>
<td>0.27*</td>
</tr>
<tr>
<td>Visual updating × necessary condition</td>
<td>−0.12</td>
</tr>
<tr>
<td>Visual updating × extraneous condition</td>
<td>−0.09</td>
</tr>
<tr>
<td>Phonological updating × visual updating × necessary condition</td>
<td>0.01</td>
</tr>
<tr>
<td>Phonological updating × visual updating × extraneous condition</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Note. Adjusted $R^2 = 0.40$, $F(11, 204) = 14.19***$  
*** $p < .001$, ** $p < .01$, * $p < .05$

Nature of problem model and updating

We required participants to execute the lexical decision task immediately after the integration process to reveal differences between problem models of high and low updating solvers. The lexical decision task for extraneous problems was analyzed.

Table 2.3 shows that the accuracy for lexical decisions was very high. Trials over the mean ±2 SD in a participant for RTs and error trials in the lexical decision task and in selecting an expression were excluded. RTs for lexical decision task were also analyzed using multiple regression analysis. To analyze target word type, we created two dummy variables: a necessary condition and an extraneous condition. Target word type was coded as follows: Necessary-information words were coded as 1 in the necessary condition and 0 in the extraneous condition. Extraneous-information words were coded as 0 in the necessary condition and 1 in the extraneous condition. Novel words were coded as 0 in the necessary condition and 0 in the extraneous condition. Two dummy variables, phonological updating, visual updating, and these interactions were used as independent variables. RTs for lexical decision task, phonological updating and visual updating were standardized. In this analysis, phonological and visual updating scores of each participant were entered repeatedly for two dummy variables. Table 2.5 shows results of this regression analysis.

The results showed significant contributions of phonological updating ($β = −0.45$, $t(204) = −4.95, p < .001$), visual updating ($β = 0.25$, $t(204) = 2.53, p < .05$), necessary condition ($β = −1.34$, $t(204) = −10.34, p < .001$), and extraneous condition ($β = −0.91$, $t(204) = −7.03, p < .001$). The results indicated that the higher the phonological updating, the shorter the RTs for lexical decisions. On the other hand, the higher the visual updating, the longer the RTs for lexical decisions. The result about necessary condition and extraneous condition indicated that RTs for
2.2. The role of updating function in constructing problem model (Experiment 2)

FIGURE 2.2: A scatter plot to reveal the relationship between the phonological updating function and RTs of lexical decision for novel words, necessary-information words, and extraneous-information words.

necessary-information words and extraneous-information words were faster than those for novel words. These results suggested that lexical decisions for necessary-information words and extraneous information words were facilitated.

Furthermore, interaction of phonological updating and extraneous condition was significant ($\beta = 0.27, t(204) = -2.08, p < .05$). This indicated that the higher phonological updating, the weaker the facilitation of lexical decision for extraneous-information words. Figure 2.2 shows a scatter plot to reveal the relationship between the phonological updating function and RTs of lexical decision for novel words, necessary-information words, and extraneous-information words.

These results suggested that both necessary- and extraneous-information words were strongly activated compared with novel words for lower phonological updating solvers. Conversely, for higher phonological updating solvers the priming effect of extraneous information was tend to be weak, while the effect of necessary information was not influenced by the phonological updating function. High phonological updating solvers formed a clear problem model by updating problem models during integration. Those with lower phonological updating, even after integration, had equal activation of necessary information and extraneous information.

In summary, results of the lexical decision task revealed that the nature of the problem model depends on individual differences related to phonological updating. Problem solvers with high phonological updating constructed a problem model that included task-relevant information only. Thus, the integration process in word problem solving depends on one’s updating function. Problem solvers with low
phonological updating might construct their problem model with all relevant and extraneous information. A problem solver’s updating function is one of the most important cognitive factors in constructing a problem model during the integration process.

2.3 General discussions about the role of updating functions in integration process

2.3.1 Constructing a Problem Model and the Role of the Updating Function

The purpose of this study was to reveal the relationship between the integration process and the updating function in solving arithmetic word problems. In two experiments, we found that integration time for an extraneous problem was longer than when necessary information alone was given. Additionally, the effect of extraneous information on integration was stronger in a low phonological updating solver than in a high phonological updating solver. These results suggested that low updating solvers struggled to form a problem model, especially for problems involving extraneous information. Therefore, the results indicate the updating function is important in the integration process in solving arithmetic word problems.

The results of Experiment 2 support the assumption that differences in the updating function cause differences in the problem model. In the lexical decision task, the higher the phonological updating, the faster the decisions. However, this effect was weaker for the extraneous-information word condition. This indicated that activation of necessary information was higher than that of extraneous information in higher phonological updating solvers. However, in lower phonological updating solvers, decisions for necessary- and extraneous-information words might be equally facilitated. This indicated the same activation of necessary and extraneous information. Note that both types of information were highly activated. High phonological updating solvers formed a problem model that included only necessary information, while low phonological updating solvers created a problem model that also included extraneous information. These problem models might influence planning time. Results of planning time showed that the higher the phonological updating, the shorter the planning time. Higher phonological updating solvers could form a correct expression based on a clear problem model, while for lower phonological updating solvers, planning should be based on an inappropriate problem model that included extraneous information.

The results of Experiment 2 are consistent with a study by Passolunghi and Pazzaglia (2004) who investigated recall error reported by low and high updating participants after solving word problems. Their results indicated that the correct recall of necessary information was higher in the high updating than in the low updating
group. High updating solvers maintained necessary information in working memory. According to our findings, these results might occur because of the activation of the information included in a problem model after the integration process.

The findings in our study revealed that the integration process in word problem solving depends on the updating function. This is because problem solvers update a problem model to form an appropriate problem model during comprehension.

### 2.3.2 Updating and Modality

Domains in the updating function should also be considered. In the present study, only phonological updating showed a relationship with integration time. Visual updating did not indicate such a relationship. The reason for this could be that processing word problems requires verbal encoding, which depends on the phonological loop. Consequently, verbal information was encoded during integration, and only updating in the phonological domain showed a relationship with integration. According to this interpretation, updating of the visuospatial sketchpad did not play an important role in solving word problems in this study. However, the visuospatial sketchpad did, in some cases, contribute more to the word problem than the phonological loop. The contribution to word problems changed with age over time even when problems were the same (Meyer et al., 2010). Second-grade students showed contribution of the phonological loop and central executive, while third-grade students demonstrated contribution of the visuospatial sketchpad. Processing word problems depended on the mental model of third-grade students who were skilled in translating words into mental representations, which requires the visuospatial sketchpad. These findings lead to a possibility that adults are more likely to encode word problems as mental models. Nevertheless, the present study did not show the importance of visual updating. Extraneous information included in the present study was of a verbal nature. Problem solvers could make decisions about relevance based on the verbal information given. The relationship between phonological updating and integration was strongly demonstrated. This suggests the possibility that problems demanding encoding of visual information could show a contribution of visual updating. Such problems (e.g., geometric problems) should be investigated to further understand the relationship between the updating function and integration.
Chapter 3

Relationship between problem format and modality of updating functions

3.1 Introduction

In Chapter 2, the importance of updating function in solving arithmetic word problems is indicated. Especially, the results suggest that the phonological updating function contribute to integration process and to construct an appropriate problem model. Arithmetic word problem solving seems to require problem solvers to manipulate phonologically coded information as part of the integration process, because they are presented using text. Previous studies have reported findings consistent with this hypothesis, namely that the phonological loop may relate to word problem solving performance (see Raghubar, Barnes, & Hecht, 2010, for a review). Andersson (2007), for example, demonstrates that phonological loop as measured by a digit span task, contributes to word problem solving performance in 10-year-old children. This suggests that phonological processing plays an important role in solving word problems. However, some researchers have reported that visual processing also relates to arithmetic word problem solving performance (Boonen, van der Schoot, van Wesel, de Vries, & Jolles, 2013; Hegarty & Kozhevnikov, 1999; Krutetskii, 1976). Hegarty and Kozhevnikov (1999), for example, report that frequency of schematic visualization of a problem relates to performance in arithmetic word problem solving. Their study indicates that visual comprehension of a problem relates to performance in solving it, which seems to show that visual manipulation is required for integration of the problem. Thus, based on these studies, there is a possibility that problem solvers use the visual updating function to integrate the information provided in an arithmetic word problem. It is still a controversial question, however, whether the phonological loop or the visuospatial sketchpad is the more important working memory component for solving word problems. Furthermore, little is known about the role of the two domains of updating, phonological and visual, in word problem solving.
3.1.1 Factors relating to the dominance of each working memory domain

There are several factors that may govern which working memory domain is dominant in word problem solving. The first plausible factor is the category of problem presented. The category of problems presented in Experiment 1 and 2 might be the reason the visual updating function has not appeared to contribute to the integration process: if the problems used require problem solvers to make particular use of phonological processing, then the visual updating function may appear not to contribute to integration. While if a word problem requires problem solvers to manipulate visual information, the visual domain could become more important. It is also possible that visual updating function is particularly important in the integration process. On the other hand, there is a possibility that individual differences in the phonological updating function always play the most important role in the integration process in word problem solving, regardless of the category of problem.

The second plausible factor is the problem format. Word problems are usually presented in a text-based format in a classroom context. It seems that phonological processing is activated for comprehension of text. This could underlie the contribution of the phonological updating function to the integration process in word problem solving. That is, there is possibility that the text-based format is an important factor in the contribution of the phonological updating function. In contrast, when a problem is presented in the form of a figure, this hypothesis suggests that problem solvers should activate visual processing in order to comprehend the problem. In this case, the visual updating function is important for problem solvers to integrate the information in the figure stimulus. If this inference is correct, there is a possibility that individual differences in the visual updating function should also be considered in the integration process in arithmetic problem solving. However, little attention has been paid to the role of the visual updating function in the integration process in word problem solving.

In this study, we were concerned with the relationship between the visual updating function and the integration process in solving of area problems presented using text or in a figure format. It is important to clarify this relationship in order to identify an effective form for the presentation of problems for children and for adults who have a weak updating function.

3.1.2 Purpose and hypothesis in this chapter

The purpose of this study was to test whether individual differences in the visual updating function are important to the integration process in word problem solving as well as phonological updating function. We used area problems as the category of problem presented. Area problems seem to require problem solvers to manipulate visual information, because they must construct a problem model based on schematic information: shape, length, etc. Area problems are commonly presented
in both text and figure formats. We investigated the relationship between the updating function in both the phonological and visual domains and the integration process in solving the same problems presented in both text and figure format.

First, we investigated whether individual differences in the visual updating function could also account for performance in the integration process when participants solve area word problems. If the category of problem explains findings of Experiment 1 and 2 that only the phonological updating function contributes to the integration process of arithmetic word problem solving, then the visual updating function could also make a contribution in solving other categories of problem. Second, we intended to investigate the relationship between the visual updating function and the integration process of solving area problems presented in the form of a figure, in order to clarify the relationships between the integration process and both domains of the updating function.

Participants in this chapter were also undergraduate students, as in Experiment 1 and 2. The contribution of working memory to arithmetic word problems is not stable in elementary school students (Meyer et al., 2010; Rasmussen & Bisanz, 2005). Moreover, if problem solvers do not have good language skills, they cannot solve arithmetic word problems, and we cannot examine the contribution of working memory to arithmetic word problem solving in such participants. Because working memory and language skills are confounded in children, we used undergraduate students with stable working memory, sufficient language skills, and relevant problem-solving schemata.

In this study, we used extraneous information to construct a condition with increased integration difficulty. Extraneous information makes the integration process more difficult, and consequently induces more errors (Muth, 1984, 1992). With undergraduate students as participants, a ceiling effect is likely for accuracy in simple arithmetic word problem solving. Therefore, accuracy rates may rarely reflect the difficulty of integration. However, extraneous information could increase the integration time required to solve arithmetic word problems. Based on previous work, we assumed that the presence of an interaction between individual difference of the updating function and the effect of extraneous information indicates a contribution by the updating function to the integration process, that is, high updating performance relates to reduce the effect of extraneous information, which increase the integration time. Therefore, we focused on reading time, especially integration time.

We hypothesized that individual differences in the visual updating function would also be important in word problem solving. In area word problem solving, the visual updating function could interact with the effect of extraneous information in their effects on integration time. Additionally, if the text format is important for the contribution of the phonological updating function in solving area word problems, we should expect visual updating to be more important in solving area problems presented in the figure format.
Furthermore, we have solved the methodological problems in analysis of Experiment 1 and 2 due to analyzing data from those experiments with multiple regression models. We repeatedly entered the updating scores of participants into the data for the factor of problem type or word condition in the analyses. This procedure can increase the probability of Type-I error because the degree of freedom can increase above the actual number of participants. Therefore, there is a possibility that the larger probability of Type-I error led to a significant \( \beta \) in these experiments. We have solved this issue by using a linear mixed-effects model regression analysis.

### 3.2 Area word problems and updating functions on phonological and visual domain (Experiment 3)

#### 3.2.1 Purpose

In Experiment 3, our purpose is to reveal the visual updating function is also important in integration process of solving area word problems. We hypothesized that problem solvers do not exhibit only the interaction effect between the problem type and the phonological updating function, but also the interaction effect between the problem type and the visual updating function in integration time of solving area word problems.

#### 3.2.2 Methods

**Participants**

Participants were 47 Japanese undergraduate students (19 males and 28 females) with an average age of 21.28 years. Most participants were recruited from an “Introduction to Psychology” lecture course; they participated voluntarily, but were compensated with course credit. Some other participants were recruited personally via acquaintances who were not particularly familiar to the researcher. All participants were native Japanese speakers and were enrolled in courses on social science, human science, engineering, natural science, or health science. Although only the engineering and natural science courses were heavily mathematical, all participants had sufficient mathematical proficiency to solve the problems presented, as demonstrated by their successful results on the National Center Test for University Admissions in Japan.

**Apparatus**

The stimulus presentation and response recording were controlled by Matlab with Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) on an Apple iMac with a 21.5-inch display and an original USB response key box with three buttons.
Task and Procedure

General procedure. Participants performed the arithmetic word problem solving task, followed by two types of updating task. The updating tasks were a letter memory task and a visual n-back task. The order of the updating tasks was counterbalanced between participants. Prior to the experiment, participants gave their written informed consent. All procedures of the study were approved by the Research Ethics Committee of the Graduate School of Humanities and Sustainable System Science at Osaka Prefecture University.

Arithmetic word problem solving task. Participants were asked to read a problem and select the formula that would provide the solution to it. The sentences making up the problem were displayed one by one. Participants could manipulate the displayed sentence, moving backward or forward by pressing the left key or middle key on the response box. They were also required to press the space key once they had comprehended the problem. The space key was enabled after the first presentation of the question statement. After this key was pressed, two expressions appeared, and participants selected the correct expression.

Following the way of Experiment 1 and 2, we collected three aspects of reading time in order to analyze the comprehension process in arithmetic word problem solving (please see Figure 2.1). The time participants took to read the sentences of a word problem, ending when the response option expressions were displayed, is referred to as the whole reading time. The time that elapsed from the onset of presentation of the first sentence of a problem to the end of the question statement’s presentation is the translation time. Finally, the time that elapsed from the question statement’s initial presentation to the onset of presentation of the response option expressions is the integration time. This last measure reflected the integration process in solving a word problem. (A fourth measure, planning time, was defined as the reaction time to select the correct expression.) We prepared two types of arithmetic word problems: necessary information and extraneous information problems. Each necessary information problem consisted of three sentences—all needed in order to form an expression that would provide the solution. Each extraneous information problem consisted of four sentences, one of which was unnecessary in forming an expression that would provide the solution.

In total, 16 area word problems were prepared for this study. They related to four two-dimensional shapes (a right-angled triangle, a triangle, a rectangle, and a parallelogram); four problems with different numerical values were composed for each shape. Half of the four problems for each shape were randomly assigned to be necessary information problems, and the other half were designated extraneous information problems.

An area word problem in this study required the information from three sentences to be solved. These sentences gave the length of the base, the height, and a question. In extraneous information problems, we added a superfluous sentence to
Chapter 3. Relationship between problem format and modality of updating functions

### Figure 3.1: Examples of problems presented (a) in text format in Experiment 3, and (b) in figure format in Experiment 4. The examples in (a) and (b) correspond to one another.

<table>
<thead>
<tr>
<th>Necessary information problem</th>
<th>Extraneous information problem</th>
<th>Length of perimeter problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>The vertical line of a rectangle is 14 cm.</td>
<td>A right-angled triangle has a vertical line of 14 cm.</td>
<td>A right-angled triangle has a vertical line of 12 cm.</td>
</tr>
<tr>
<td>The horizontal line is 20 cm.</td>
<td>The length of the horizontal line is 9 cm.</td>
<td>The length of the horizontal line is 7 cm.</td>
</tr>
<tr>
<td>What is the area of this rectangle?</td>
<td>The length of the hypotenuse is 17 cm.</td>
<td>The length of the hypotenuse is 14 cm.</td>
</tr>
<tr>
<td></td>
<td>What is the area of this triangle?</td>
<td>How long is the perimeter of this triangle?</td>
</tr>
</tbody>
</table>

![Figure 3.1](image-url)
these sentences, giving the length of an oblique or diagonal line. The position of the extraneous information sentence (as the second or third sentence) was randomized. Examples of area problems (a necessary information and an extraneous information problem) are presented in Figure 3.1a.

If the experiment consisted solely of area problems, participants might completely ignore the extraneous information sentences. To avoid this, we prepared four problems which asked the participant to calculate the perimeter of a shape. To solve perimeter problems, it is necessary to use the length of an oblique or diagonal line, which was the extraneous information provided in the area problems. Two perimeter problems were prepared for each of the triangle and parallelogram shapes. An example of perimeter problem is also presented in Figure 3.1a. In the same way, 10 word problems were prepared for practice trials. On practice trials, the first sentence was presented along with a figure (without numerical values) in order to activate the schema of an area problem. This figure disappeared after participants first pressed the forward key, and did not reappear if they returned to the first sentence of the problem. These figures were not presented in experimental trials.

**Updating tasks.** Updating tasks in this experiment were adapted from Experiment 1 and 2 and Agostino et al. (2010). A letter memory task measured phonological updating in working memory. In this task, letters were presented in sequence for 2000 ms per letter on a computer screen. Participants were required to rehearse aloud the last four letters throughout, by dropping the fifth most recent letter and adding the newest. There were four stimulus sets, consisting of 5, 7, 9, and 11 letters. The participant’s phonological updating score was their percentage of correct responses.

A visual n-back task measured visual updating in working memory. On each trial, 10 squares, one of which was black and the others white, were presented for 500 ms on a computer screen. A white circle was then presented at the center of the screen. During presentation of the white circle, participants were required to press a key indicating whether the pattern just displayed was the same as the one displayed n trials previously. The white circle changed to a black circle after a key was pressed. The duration of circle presentation was 1500 ms. Figure 3.2 presents examples of this stimulus sequence in the 2-back task. In each n-back task, there were 12 practice trials followed by 54 test trials. Twenty test trials were designated as matching, and 34 as mismatching. We employed 2-back and 3-back tasks. To avoid a response bias in scoring, each participant’s score on each task was calculated using the formula (proportion of correct matches + proportion of correct mismatches) / 2. This calculation ensured that match trials and mismatch trials were weighted equally.

### 3.2.3 Results and discussions

The number of movements between sentences while solving word problems across all participants had a mean of 97.68 and standard deviation (SD) of 32.58. We treated
Chapter 3. Relationship between problem format and modality of updating functions

Fixation displayed for 500 ms.
Stimulus displayed for 1000 ms.
Response required within 1500 ms.
Response cue: white circle.
Response acknowledgement: black circle.

In the absence of a response, the white circle remained on-screen.

The correct response was “Mismatch”.

The correct response was “Match”.

FIGURE 3.2: An example of a display sequence in the visual 2-back task. Participants determined whether the current stimulus matched the 2-back stimulus. (In the 3-back task, they determined whether the current stimulus matched the 3-back stimulus.) After each stimulus (except the first n), a white circle was presented. Participants were required to respond during presentation of this cue. A black circle was presented after the participant’s response in order to indicate that the response had been recorded.
3.2. Area word problems and updating functions on phonological and visual domain (Experiment 3)

<table>
<thead>
<tr>
<th>Task</th>
<th>M</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonological updating</td>
<td>0.65</td>
<td>0.23</td>
<td>0.19</td>
<td>1.00</td>
</tr>
<tr>
<td>Visual updating</td>
<td>0.84</td>
<td>0.13</td>
<td>0.45</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Note. The phonological updating and visual updating values are given as rates.

Table 3.2: Means and standard deviations (in parentheses) for whole reading time, translation time, integration time, planning time, and accuracy by problem type in Experiment 3.

<table>
<thead>
<tr>
<th>Necessary information</th>
<th>Extraneous information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole reading time</td>
<td>7412.38 (2203.79)</td>
</tr>
<tr>
<td>Translation time</td>
<td>7290.62 (2152.29)</td>
</tr>
<tr>
<td>Integration time</td>
<td>1737.21 (613.14)</td>
</tr>
<tr>
<td>Planning time</td>
<td>926.69 (331.94)</td>
</tr>
<tr>
<td>Accuracy</td>
<td>1.00 (0.03)</td>
</tr>
</tbody>
</table>

Note. Whole reading time, translation time, integration time, and planning time are given in milliseconds. Accuracies are given as rates.

participants as outliers if their number of movements fell above the mean + 3 $SD$. One participant was excluded from the analysis on this criterion; subsequent analyses were thus conducted for 46 participants.

Whole reading, translation, integration, and planning times were calculated after excluding perimeter problem trials and erroneous trials, on which participants selected the wrong expression as the answer to the arithmetic word problem. Trials were also excluded if an individual’s whole reading time fell outside the mean ±2 $SD$ in the participant. Descriptive statistics for updating function measures are presented in Table 3.1, and descriptive statistics for whole reading, translation, integration, and planning times and accuracy are presented in Table 3.2.

Participants’ accuracy at selecting the correct expression was analyzed using a paired-samples t-test of word problem type (necessary information problems vs. extraneous information problems). The difference in accuracy between the necessary information problems and extraneous information problems was significant ($t(45) = 2.88, p < .05$). This result indicates that participants gave more wrong answers to extraneous than to necessary information problems, meaning the former were more difficult than the latter to solve.

We constructed linear mixed-effects regression models to estimate the contribution of the updating functions in solving word problems. A linear mixed-effects regression model resolves an issue that a probability of Type-I error was increased in analyses of Experiment 1 and 2. Whole reading, translation, integration, and planning times were used as dependent variables. Problem type, phonological updating
Chapter 3. Relationship between problem format and modality of updating functions

Score, visual updating score, and their interactions were entered as fixed-effects variables. A random-effect variable of participant was also included. Prior to analysis, participants’ phonological and visual updating scores were centered on the mean. Problem type was coded as a dummy variable, with 0 representing a necessary information problem and 1 an extraneous information problem. Table 3.3 presents the results of these linear mixed-effects regression model analyses.

Analysis of whole reading times and translation times showed a significant effect of problem type ($\beta = 2894.54, t(633) = 15.29, p < .001; \beta = 2357.72, t(633) = 12.68, p < .001$). This result indicates that problem solvers required more time to solve an extraneous information problem than a necessary information problem, because they had to read an additional sentence in the extraneous information problem.

Analysis of integration times also showed a significant effect of problem type ($\beta = 1000.27, t(633) = 10.74, p < .001$). Furthermore, this analysis revealed a significant three-way interaction between problem type, phonological updating score, and visual updating score ($\beta = 7089.25, t(633) = 2.43, p < .05$). To understand this interaction in more detail, we conducted slope-difference tests (Dawson & Richter, 2006). First, we computed simple slopes for problem type at combinations of high and low values of the phonological and visual updating scores. High and low values were calculated by shifting the scores by $\pm 1$ SD (Figure 3.3). This meant that the slope indicated the effect of extraneous information. To determine whether the difference between a pair of slopes was significant, we tested whether the ratio of the slope difference and its standard error differed from zero. This analysis showed a significant difference between the slopes at high phonological and low visual updating scores, and between the slopes at low phonological and low visual updating scores ($t(675) = 2.17, p < .05$). This analysis indicated that the slope was shallower at high phonological and low visual updating scores than at low phonological and low visual updating scores. The effect of extraneous information could be reduced by strong phonological updating. Furthermore, the slope difference test showed that the slope at low phonological and high visual updating scores was significantly shallower than the slope at low phonological and low visual updating scores ($t(675) = -2.49, p < .05$). This result suggests that a strong visual updating function contributed to a reduction in the effect of extraneous information, which increased integration time. These results indicate that, in addition to the phonological updating function, the visual updating function also interacted with problem type in the integration process. This supports our hypothesis that the visual updating function is also important for solving arithmetic word problem when the category of problem requires problem solvers to manipulate visual information.
### Table 3.3: Coefficient estimates in linear mixed-effects regression models of whole reading time, translation time, integration time, and planning time in Experiment 3.

<table>
<thead>
<tr>
<th></th>
<th>Whole reading time $\hat{\beta}$</th>
<th>Translation time $\hat{\beta}$</th>
<th>Integration time $\hat{\beta}$</th>
<th>Planning time $\hat{\beta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem type</td>
<td>2894.54** [2522.78, 3266.30]</td>
<td>2357.72** [1992.71, 2722.73]</td>
<td>1000.27** [817.44, 1183.09]</td>
<td>325.90** [172.60, 479.21]</td>
</tr>
<tr>
<td>Phonological updating</td>
<td>−93.66 [−3200.52, 3013.20]</td>
<td>−153.52 [−3141.13, 2834.09]</td>
<td>102.07 [−944.43, 1148.57]</td>
<td>−130.60 [−764.52, 503.31]</td>
</tr>
<tr>
<td>Problem type × phonological updating</td>
<td>−1113.59 [−2745.37, 518.19]</td>
<td>−1306.78 [−2908.95, 295.38]</td>
<td>4.02 [−798.47, 806.50]</td>
<td>160.43 [−512.47, 833.33]</td>
</tr>
<tr>
<td>Problem type × visual updating</td>
<td>−658.70 [−3760.44, 2443.04]</td>
<td>−566.84 [−3612.29, 2478.61]</td>
<td>−539.80 [−2065.22, 985.62]</td>
<td>−599.78 [−679.38, 1878.94]</td>
</tr>
<tr>
<td>Phonological updating × visual updating</td>
<td>−1307.44 [−23 318.08, 20 703.20]</td>
<td>−71.43 [−21 239.91, 21 097.05]</td>
<td>−1600.16 [−9038.93, 5838.62]</td>
<td>−452.03 [−4983.90, 4079.84]</td>
</tr>
<tr>
<td>Three-way interaction</td>
<td>7954.43 [−3655.81, 19 564.67]</td>
<td>2569.91 [−8829.66, 13 969.48]</td>
<td>7089.25* [1379.18, 12 799.32]</td>
<td>1760.01 [−3028.65, 6548.66]</td>
</tr>
<tr>
<td>Constant</td>
<td>7417.85** [6724.83, 8110.86]</td>
<td>7294.32** [6627.91, 7960.72]</td>
<td>1742.78** [1509.42, 1976.15]</td>
<td>930.49** [789.20, 1071.78]</td>
</tr>
</tbody>
</table>

**Note.** Confidence intervals are indicated in brackets: [lower limit, upper limit]. **$p < .01$, *$p < .05$, +$p < .10$.**
The difference between the slope for problem type at high phonological and low visual updating scores and at low phonological and high visual updating scores was not significant ($t(675) = 0.054, n.s.$). The difference between the slope for problem type at high phonological and high visual updating scores and at low phonological and low visual updating scores was also non-significant ($t(675) = -0.529, n.s.$). This suggests that there was no reduction in the effect of extraneous information in problem solvers who had strong phonological and strong visual updating functions.

Analysis of planning time showed a significant effect of problem type ($\beta = 325.90, t(633) = 4.17, p < .001$). This indicates that participants required more time to decide on the correct expression for an extraneous information problem than for a necessary information problem.

Post hoc power analyses for each of our models (of whole reading time, translation time, integration time, and planning time) were conducted. At $\alpha = 5\%$, statistical power of 100% was achieved for whole reading time ($f^2 = 0.24$), translation time ($f^2 = 0.18$), and integration time ($f^2 = 0.18$), while power of 89% was achieved for planning time ($f^2 = 0.026$). Because statistical power exceeded 80% in every case (although the power or effect size for planning time was lower than that of the other dependent variables), it seems that this experiment had sufficient power to identify the effects of all these models.

In summary, the results of this experiment indicate that the phonological updating function contributes to the process of integration in area word problem solving, according to the outcome of the slope-difference test. This result supports the findings of Experiment 1 and 2, in which the phonological updating function played an important role in the integration process. Furthermore, Experiment 3 suggests that the visual updating function also contributes to integration, because the slope for problem type was shallower at low phonological and high visual updating scores than at low phonological and low visual updating scores. Experiments 1 and 2 shows that phonological updating uniquely contributes to the integration process in solving an arithmetic word problem. Results in this experiment, however, suggest that both phonological and visual updating contribute to integration in an area word problem. The difference between the previous experiments and this experiment can be explained by the category of problem presented. This experiment used area problems, which could load the visual domain. Therefore, results in Experiment 3 supported our hypothesis that the visual updating function is also important in arithmetic word problem solving when the category of problem requires problem solvers to manipulate visual information.

However, there was also a complicating finding: the slope for problem type at high phonological and high visual updating scores did not differ from the slope at low phonological and low visual updating scores. This suggests that a strong updating function did not contribute to a reduction in the effect of extraneous information when the updating functions in both domains were strong. Because the results of slope difference tests suggested that a strong updating function in a single domain,
3.2. Area word problems and updating functions on phonological and visual domain (Experiment 3)

**Figure 3.3:** Visualization of the three-way interaction between problem type, phonological updating score, and visual updating score. Each slope was calculated by shifting the updating score values by ±1 SD. For example, the slope of problem type at a high phonological updating and low visual updating score was the slope of the model when 1 SD was added to the phonological updating score and 1 SD was subtracted from the visual updating score. Red lines with asterisks indicate pairs of significantly different slopes.
phonological or visual, reduced the effect of extraneous information, it was expected
that this effect would also be reduced when both the phonological and visual updat-
ing functions were strong. This hypothesis was not supported by our results. We
were unable to ascertain why the experiment produced this contradictory finding.
Further investigation is needed to understand in greater detail the processing of area
word problems in problem solvers with strong updating functions in both domains.

In Experiment 3, area word problems were presented in a textual format. It could
be argued that the contribution of the phonological updating function arose from the
problem format, while that of the visual updating function arose from the category
of problem. On the other hand, there is possibility that attempting to solve an area
problem activates both the phonological and the visual updating functions regard-
less of the problem format. The relationship between the importance of each updat-
ing domain and problem format is unclear. Further investigations were needed to
reveal which of two possibilities is more appropriately.

3.3 Area problems in a figure format and updating functions
on phonological and visual domain (Experiment 4)

3.3.1 Purpose

In Experiment 4, our purpose is to reveal the relationship between the importance of
each updating domain and problem format. We hypothesized that visual updating
function would uniquely contribute to the integration process in solving area prob-
lems presented in a figure format. If problem format influences the importance of
the phonological or visual updating functions, the results could show a significant
interaction between problem type and visual updating score, but not phonological
updating score.

3.3.2 Methods

Participants

Participants were 49 undergraduate students (21 males and 28 females) in Japan
with an average age of 19.63 years. Most participants were recruited from an “Intro-
duction to Psychology” lecture course; they participated voluntarily but were com-
penated with an honorarium of 1000 yen. All participants were native Japanese
speakers and were enrolled in courses in social science, human science, engineer-
ing, natural science, or health science. Although only the engineering and natural
science courses were heavily mathematical, participants on other courses had suffi-
cient mathematical proficiency to solve the problems presented, as demonstrated by
their successful results on the National Center Test for University Admissions.
Task and Procedure

The same procedures were used as in Experiment 3, but the area problems were presented in a figure format. The figure included the shape and the length of a side or diagonal. Participants were instructed to select the correct expression to calculate the area of the shape when a question mark was presented inside the figure, and to select the correct expression to calculate the perimeter of the shape when a red line and question mark were presented around the figure. This flow of presentation corresponded with the flow of the sentences presented in Experiment 3. Figure 3.1b shows examples of the figures presented in Experiment 4.

Participants solved 20 arithmetic problems (following 10 practice problems), then completed the updating tasks. The order of the updating tasks (i.e., whether the first task was the phonological or the visual updating task) was counterbalanced between participants. In Experiment 3, participants’ visual updating function scores were calculated only on the basis of their performance in the 3-back task. Thus, in order to reduce the load of the experiment for participants, we dropped the 2-back task in Experiment 4 and used only the 3-back task. The procedures used in Experiment 4 were approved by the Research Ethics Committee of the Graduate School of Humanities and Sustainable System Science at Osaka Prefecture University.

3.3.3 Results and Discussion

The same procedures used in Experiment 3 were followed for the analysis of the data from Experiment 4. Descriptive statistics for updating function scores are presented in Table 3.4, and descriptive statistics for whole reading, translation, integration, and planning times and accuracy are presented in Table 3.5. Participants’ accuracies in selecting the correct expression were analyzed using a paired-samples t-test of problem type (necessary information problem vs. extraneous information problem). The difference in accuracy between the necessary information problems and extraneous information problems was not significant ($t(48) = 0.94, n.s.$). This result indicates that the difficulty of selecting the correct expression was the same for both types of problem in Experiment 4.

In the same way as for Experiment 3, we constructed linear mixed-effects regression models to estimate the contribution of the updating functions in solving word problems. Table 3.6 gives the results of these linear mixed-effects regression models. Analysis of whole reading time showed a significant effect of problem type ($\beta = 2219.75, t(691) = 18.76, p < .001$): a longer whole reading time was required to solve an extraneous information problem. This analysis also showed a significant interaction between problem type and visual updating score ($\beta = -1435.91, t(691) = -2.02, p < .05$). This interaction indicates that the higher the visual updating function score, the smaller the increase in whole reading time produced by extraneous information.
Chapter 3. Relationship between problem format and modality of updating functions

Table 3.4: Descriptive statistics for scores on the phonological and visual updating tasks in Experiment 4.

<table>
<thead>
<tr>
<th>Task</th>
<th>M</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonological updating</td>
<td>0.65</td>
<td>0.22</td>
<td>0</td>
<td>0.94</td>
</tr>
<tr>
<td>Visual updating</td>
<td>0.72</td>
<td>0.18</td>
<td>0.24</td>
<td>0.98</td>
</tr>
</tbody>
</table>

*Note.* The phonological updating and visual updating values are given as rates.

Table 3.5: Means and standard deviations (in parentheses) for whole reading time, translation time, integration time, planning time, and accuracy by problem type in Experiment 4

<table>
<thead>
<tr>
<th></th>
<th>Necessary information</th>
<th>Extraneous information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole reading time</td>
<td>3719.85 (1217.45)</td>
<td>5983.09 (2011.36)</td>
</tr>
<tr>
<td>Translation time</td>
<td>3536.95 (1290.41)</td>
<td>5447.49 (2144.36)</td>
</tr>
<tr>
<td>Integration time</td>
<td>1488.78 (541.05)</td>
<td>2126.89 (731.27)</td>
</tr>
<tr>
<td>Planning time</td>
<td>1059.10 (308.10)</td>
<td>1309.97 (508.12)</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.98 (0.08)</td>
<td>0.97 (0.08)</td>
</tr>
</tbody>
</table>

*Note.* Whole reading time, translation time, integration time, and planning time are given in milliseconds. Accuracies are given as rates.

Analyses of translation and integration times each showed a significant effect of problem type ($\beta = 1868.41, t(691) = 16.48, p < .001; \beta = 641.52, t(691) = 8.50, p < .001$): participants needed longer to complete the translation and integration processes when solving an extraneous information problem than when solving a necessary information problem. Analysis of planning time also showed a significant effect of problem type ($\beta = 325.90, t(691) = 3.72, p < .001$). Participants required more time to decide on the correct expression for an extraneous information problem than for a necessary information problem.

Post hoc power analyses were conducted for each dependent variable. At $\alpha = 5\%$, power of 100% was achieved for whole reading time ($f^2 = 0.35$), translation time ($f^2 = 0.27$), and integration time ($f^2 = 0.09$), while power of 98% was achieved for planning time ($f^2 = 0.03$). As power in this experiment exceeded 80% in every case, it seems that this experiment provided sufficient power to identify the effects of model.
TABLE 3.6: Coefficient estimates in linear mixed-effects regression models of whole reading time, translation time, integration time, and planning time in Experiment 4.

<table>
<thead>
<tr>
<th></th>
<th>Whole reading time $\hat{\beta}$</th>
<th>Translation time $\hat{\beta}$</th>
<th>Integration time $\hat{\beta}$</th>
<th>Planning time $\hat{\beta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem type</td>
<td>2219.75$^{**}$ [1987.48, 2452.02]</td>
<td>1868.41$^{**}$ [1645.70, 2091.12]</td>
<td>641.52$^{**}$ [493.18, 789.87]</td>
<td>325.90$^{**}$ [115.04, 379.57]</td>
</tr>
<tr>
<td>Phonological updating</td>
<td>-767.61 [-2848.30, 1313.09]</td>
<td>-226.91 [-2436.82, 1983]</td>
<td>-438.75 [-1288.27, 410.76]</td>
<td>-130.60 [-909.41, 182.62]</td>
</tr>
<tr>
<td>Visual updating</td>
<td>-1713.20 [-4433.74, 1007.34]</td>
<td>-2258.36 [-5146.10, 629.38]</td>
<td>-423.44 [-1539.22, 692.34]</td>
<td>-369.56 [-1280.32, 162.91]</td>
</tr>
<tr>
<td>Problem type × phonological updating</td>
<td>-29.49 [-1073.70, 1014.72]</td>
<td>-77.51 [-2105.27, 556.12]</td>
<td>541.90 [-125.05, 1208.85]</td>
<td>160.43 [-291.43, 1288.21]</td>
</tr>
<tr>
<td>Problem type × visual updating</td>
<td>-1435.91$^{*}$ [-2823.70, -48.12]</td>
<td>-774.57 [-2105.27, 556.12]</td>
<td>-401.86 [-1288.03, 484.31]</td>
<td>599.78 [-291.43, 1288.21]</td>
</tr>
<tr>
<td>Phonological updating × visual updating</td>
<td>2903.78 [-11 289.48, 17 097.04]</td>
<td>5445.34 [-9619.12, 20 509.81]</td>
<td>-1220.84 [-7045.06, 4603.38]</td>
<td>-452.03 [-2668.19, 4870.44]</td>
</tr>
<tr>
<td>Three-way interaction</td>
<td>3235.37 [-3975.48, 10 446.21]</td>
<td>2685.49 [-4228.68, 9599.66]</td>
<td>-1178.93 [-5783.58, 3425.72]</td>
<td>1760.01 [-4412.57, 3796.05]</td>
</tr>
<tr>
<td>Constant</td>
<td>3697.05$^{**}$ [3246.97, 4147.12]</td>
<td>3494.32$^{**}$ [3016.36, 3972.27]</td>
<td>1500.02$^{**}$ [1316.04, 1684]</td>
<td>930.49$^{**}$ [937.14, 1174.08]</td>
</tr>
</tbody>
</table>

Note. Confidence intervals are indicated in brackets: [lower limit, upper limit]. $^{**}p < .01$, $^{*}p < .05$, $^{+}p < .10$. 

and visual domain (Experiment 4)

3.3. Area problems in a figure format and updating functions on phonological
We hypothesized that the interaction between problem type and updating functions observed in Experiment 3 for integration time would be replicated in Experiment 4; however, this interaction was observed in Experiment 4 only for whole reading time, and not for integration time. It is necessary to discuss this discrepancy in order to understand the role of the updating function in area problem solving. One explanation for this contradictory result may be that the problem-solving process was affected by the problem format (textual format or figure format). Whole reading times were shorter in Experiment 4 than in Experiment 3. This indicates that the figure format enabled problem solvers to complete the translation and integration processes relatively quickly. In particular, translation time was shorter in Experiment 4 (4492 ms) than in Experiment 3 (8474 ms). This finding suggests that the translation load was reduced by presenting problems in the figure format. In addition, integration time was also shorter in Experiment 4 (1808 ms) than in Experiment 3 (2244 ms). However, there is no apparent rationale for why the integration process required less time in Experiment 4 than in Experiment 3. Even if the problem format was different, problem solvers still had to integrate some information which they translated into their working memory. Shorter integration time suggests that the integration process was included outside the integration time.

The reduced translation and integration time for figure format problems in the present study suggest that problem solvers were able to integrate the information in a problem in figure format during its first presentation. In support of this, Orrantia, Múñez, and Tarín (2014) suggest that problem solvers may construct a temporal mental model before reading all sentences of a problem when it is presented in text. Thus, the integration process in Experiment 4 could be distributed between the periods regarded as translation time and integration time. For these reasons, it seems that the interaction between problem type and visual updating score was reflected in the whole reading time rather than in the integration time specifically.

Therefore, we interpret these results as suggesting that the visual updating function contributed to the integration process in problem solving in this experiment: that is, a strong visual updating function reduced the effect of extraneous information. In contrast, no interaction between problem type and the phonological updating function was observed. This suggests that the phonological updating function did not contribute to the process of solving area problems presented in a figure format. These findings support our hypothesis that only individual differences in the visual updating function are important for problem solving when both the category and the format of the problem require problem solvers to manipulate visual information.
3.4 General discussions about the modality of updating function in integration of problem solving

The aim of Experiment 3 and 4 was to test whether the visual updating function plays an important role in arithmetic problem solving. Experiment 3 investigated whether the visual updating function plays an important role in integration of solving area word problems. The effect of extraneous information, which increases integration difficulty, was found to be reduced for problem solvers with a strong phonological or visual updating function. These results suggest that, in solving an area word problem, both the phonological and visual updating functions play an important role in integrating the information presented in the problem.

Experiment 4 was conducted to investigate the role of individual differences in the visual updating function in solving area problems presented in a figure format. The results of Experiment 4 showed that individual differences in the visual updating function have an important effect on solving time for area problems presented in a figure format.

Taken together, the results of the two experiments suggest that individual differences in the visual updating function have an important effect on solving arithmetic problems when the category of problem or the problem format requires more visual processing. In other words, the category and format of a problem affect which domain of updating function is dominant in solving it. The domain of the updating function used for problem solving is constrained by the category of problem or problem format: it is not always that the problem solvers could opt to use their stronger updating domain. For example, the results of Experiment 4 suggest that even if a problem solver had a strong phonological updating function, their proficient domain (i.e., the phonological domain) could not contribute effectively to solving area problems presented in a figure format.

However, the results of Experiment 3 suggest that problem solvers select which domain of updating they will use to integrate a word problem. An area word problem seems to require both phonological and visual processing, because the problem is presented in text and describes visual information about a shape. Problem solvers who have a stronger phonological updating function can integrate the information in an area word problem using the phonological domain; conversely, problem solvers who have a stronger visual updating function can do so using the visual domain. In relation to this interpretation, there is further evidence that which domain is used for solving depends on the problem solver. Meyer et al. (2010) demonstrate that performance on a phonological loop task predicts performance on word problem solving in second-graders; in contrast, in third-graders, word problem solving performance is predicted not by phonological loop performance but by visuospatial sketchpad function. The authors argue that second-graders use phonologically coded information and third-graders visually coded information to solve the problems. Their study shows that even if problem solvers are presented with the same
word problems, the important working memory domain could differ depending on the problem solver. Given those results, it is possible that the findings of Experiment 3 imply that even solving the same word problems, problem solvers will use their respective stronger updating functions in the integration process.

Further investigation is needed to understand the relationship between our findings and individual processing strategies. For example, Krutetskii (1976) argues that there are three types of problem solver, categorized according to how they process mathematical information. The first is the analytic type, who prefer verbal rather than visual processing in problem solving; the second is the geometric type, who prefer to use visual imagery; and the third is the harmonic type, who have no tendency one way or the other. Furthermore, Hegarty and Kozhevnikov (1999) demonstrate two types of visual spatial strategy, namely a schematic representation strategy and a pictorial representation strategy. Understanding the relationship between proficiency in the two domains of updating and these strategies seems to be important, because the interaction between strategy and the more proficient (phonological or visual) domain of updating could play an important role in integrating the information presented in a problem.

Based on our results, we infer that the category of problem, the problem format, and also the problem solver’s strategy could influence the importance of the role played by updating functions in arithmetic problem solving. There are complex relationships between these factors. Our findings in Experiment 3 suggested that problem solvers who had both high phonological and high visual updating scores could not utilize their strong updating ability to reduce the effect of extraneous information. It could be argued that a strong updating function is not an advantage in all situations. Similarly, even if problem solvers had a weak phonological updating function, those who had a strong visual updating function could integrate the information contained in a word problem without difficulty for some categories of problem or problem formats. Our study suggests that we should consider the interactions between these factors in order to understand the role of individual differences in the integration process of solving arithmetic problems.
Chapter 4

General discussion

4.1 The updating function and integration process in arithmetic word problem solving

Although the integration process is thought to be the most important process in word problem solving, the underlying cognitive functions have not been made clear thus far. We focus on the updating function, one of the executive functions in working memory, as an important cognitive function in the integration of information during arithmetic word problem solving. In Experiment 1, we showed that the updating function plays an important role in the integration process. The results of Experiment 1 indicate that the effect of extraneous information on integration time was lower when the score of the updating function was higher. Although extraneous information increases the difficulty of the integration process, problem solvers who have a high updating function might not be susceptible to the influence of the extraneous information. The way that the updating function contributes to the integration process was revealed in Experiment 2, which investigated the state of information in the problem model immediately after integration. To examine the state of information, a priming paradigm was used. We investigated whether the size of the priming effect differs for necessary-information words or extraneous-information words. It was assumed that highly applicable information in the problem model is determined rapidly in the lexical decision task because the information was highly activated. The higher updating function was thus likely to exhibit a lower priming effect for extraneous-information words. In other words, problem solvers with lower updating functions were likelier to make a problem model that included the extraneous information. These findings indicate that the updating function plays a role in re-constructing the problem model in order to make a more appropriate problem model during the integration process.

In Experiments 1 and 2, only the score measured on the phonological updating task exhibited a relationship with the integration process. However, the results of Experiment 3 indicate that the visual updating function also plays an important role in integration while solving a word problem, at least when the word problem asks problem solvers to determine the area of a shape. If a problem-solving task was
presented in a figure format, only the visual updating function exhibited a contribution to the integration process in solving the area problem (in Experiment 4). These findings imply that the respective importance of the phonological or visual updating function depends on the category of the problem and the format of the problem presentation. In solving area problems, the fact that the phonological and visual updating functions each contribute to integration suggests that some problem solvers use the phonological updating function more for integration, while the others rely on the visual updating function. According to this finding, there is the possibility that a problem solver’s strategy also influences the importance of the phonological and visual updating functions.

Previous studies suggested that integrating problem information is important to success in solving arithmetic word problems. The relevant information in the problem should be held in the problem solver’s working memory, and either the phonological loop or the visuospatial sketchpad could play an important role in holding the information. The information held in working memory might be manipulated in order to construct an appropriate problem model in the integration process via the central executive function. Our findings in this study indicate that the updating function contributes to the manipulation required to construct the problem model. Where previous studies did not demonstrate the role of the updating function in arithmetic word problem solving, our study found that the updating function plays an important role in the integration process when constructing a problem model. Furthermore, there is a need to consider both the phonological and visual updating functions. Problem solvers could translate problem sentences into a phonological or visuospatial code, the updating function could manipulate the information in the phonological loop or visuospatial sketchpad. Indeed, the phonological and visual updating functions exhibit the different contributions depending on the category or format of the problem.

It is important to consider whether the facilitation of integration was an effect of updating or simply working memory capacity (WMC). It may be the case that an individual with a large WMC can perform updating tasks efficiently and reduce integration time because he or she can store all information in working memory (Daneman & Carpenter, 1980). If this is the case, such individuals may be expected to integrate all the information easily into a problem model. However, this possibility was rejected based on our results from Experiment 2 (in Chapter 2). The results of the lexical decision task indicate that the facilitation effects weakened for extraneous-information words, as the level of the phonological updating function increased. This means that extraneous information was held in working memory even if problem solver’s updating function was low. This was inconsistent with the account of simple WMC because the capacity of low phonological updating individuals sufficed to store all the information. As high updating solvers updated their problem model, these results were more likely due to the updating function than WMC.
4.1. The updating function and integration process in arithmetic word problem solving

On the other hand, in recent studies, WMC seemed to exhibit the ability to control our attention (e.g., Engle, Tuholski, Laughlin, & Conway, 1999; Kane, Bleckley, Conway, & Engle, 2001). Controlled attention is referred to an executive control capability; that is, the ability to effectively maintain stimulus, goal, or context information in an active, easily accessible state in the face of interference, to effectively inhibit goal-irrelevant stimuli or responses, or both (Kane et al., 2001). For example, high-WMC individuals were able to actively maintain information in memory while simultaneously attending to a secondary-processing task. From this point of view, the present study’s results reflected general attentional control underlying the WMC rather than the updating function. Indeed, the updating function has been shown to be highly correlated with WMC as measured by the operation span task (Miyake et al., 2000). However, dual-task performance did not relate to the updating function (Miyake et al., 2000), which related to the WMC (Piai & Roelofs, 2013). Although there is a possibility that the updating function is an underlying process of WMC, there is controversy over the relationship between the updating function and WMC. There is a similar controversy over the relationship between the inhibition function and the updating function. Some might argue that the results of this study could be interpreted as showing that extraneous information, which is task-irrelevant information, is inhibited. However, Miyake et al. (2000) shows that the inhibition function and updating function are separable functions because they do not show the same relationship with other complex task performances. For example, updating relates to the WMC, while inhibition dose not. Our results suggest that in constructing an appropriate problem model, problem solvers activate necessary information and deactivate extraneous information held in working memory at least once. That is, problem solvers could need to change an older problem model to a newer one in their working memory. This view of the updating function seems to better account for our results, although further investigation is needed to solve the above problems.

Furthermore, some might argue that the relationship between the score on the updating function task and integration time, especially in Experiments 3 and 4, indicates the importance of short-term storage but not the updating function. In Chapter 3, the interaction between the strengths in the different domains of the updating function and problem type differed for area problems presented textually or in the form of figures. Our findings suggest that area word problems recruit both phonological and visual updating, while area problems presented in a figure format recruit only visual updating. However, our results might be explained by the capacity of the phonological loop or visuospatial sketchpad. Short-term storage, which is independent of the updating function, could contribute to problem solving; that is, phonological storage may be important when problems are likely to require phonological processing, while visual storage may be important when visual processing is required by the problem (as in the figure format). Although our findings suggest
that the performance of the updating function interacted with the degree of integration difficulty presented by the presence of extraneous information, it is unclear whether this relationship was truly due to the proficiency of the updating function, or whether it could be attributed to participants’ proficiency in tapping short-term storage during tasks involving the updating function. This uncertainty is a form of the task-impurity problem (Miyake et al., 2000). Our experiment did not extract the proficiency of the updating function as a pure latent variable using a structural equation model, or compare the updating scores with the scores on a short-term memory task. For this reason, it could be argued that our results simply revealed the relationship between the proficiencies of the two types of short-term storage and arithmetic problem solving.

It is certainly possible that the updating task could partially tap into the proficiency of short-term storage. Our study revealed that the visual updating function is more important for problems presented in the figure format than those presented in the text format. However, an interpretation based solely on the simple contribution of short-term storage cannot explain our findings. To exhibit a reduced effect of extraneous information, problem solvers must handle the difficulty this presents for the integration process; to do so, they are required to maintain information but also perform some processing. Indeed, an interaction between the updating function score and problem type was observed in this study; that is, the influence of extraneous information on integration difficulty depended on individual differences in updating task performance. Additionally, this interaction was found in integration times in Experiment 3 (also Experiment 1 and 2). This finding supports the interpretation that the interaction between the updating task score and problem type was not due to simple short-term storage but to the updating function itself. We therefore think it more appropriate to conclude from our results that individual differences in the strength of the updating function are important to the integration process in arithmetic word problem solving.

However, the category of problem or problem format actually influenced the importance of the phonological and visual updating functions. This implies the independent influence of each storage system. We observed individual differences as measured by tasks requiring participants to update information in their working memory for particular types of storage (phonological or visual). According to our findings, the individual differences in updating function, and in particular the matter of which problem solvers update information in a given storage, might be important in solving arithmetic word problems. It is unclear that a pure updating function is important in solving an arithmetic word problem, or the ability to update information in a specific domain is important. An investigation is necessary that measures the capacities of each short-term storage and the pure updating function as latent variables. Further investigation is needed to reveal the roles of each short-term storage, the pure updating function, and the specific updating function in each domain in the integration process when solving arithmetic word problems.
4.2 Educational implications

Although the present study investigated a particular cognitive process required in arithmetic word problem solving, which is a type of problem encountered in school, the participants were not elementary schoolchildren. Therefore, further investigation of the topic in elementary schoolchildren is needed. However, our findings imply that individual differences in the updating function in both domains play an important role in the integration process, and that the presentation format of the problem governs the importance of individual differences in the updating function. Our results suggest that performance in word problem solving depends on the relationship between the updating function in each domain and the requirements of the problem at hand. Further research should investigate the relationships among the category of problem, the problem format, and the updating function in children who struggle to solve word problems.

However, there are educational implications that may be drawn from this study. Our findings suggest that less able updating function causes unsuccessful integration. This is an explanation about unsuccessful integration from a cognitive process perspective. Our findings suggest the possibility that children with lower updating have a problem with the integration process. These children might improve their performance in solving word problems if we could train their updating function with working memory training. However, some studies report short-term improvements of working memory training only for working-memory-related tasks, while other general skills (e.g., arithmetic, academic skills, and verbal ability) are not improved (Melby-Lervåg & Hulme, 2013; Sala & Gobet, 2017). On the other hand, Au et al. (2015) reports that working memory training, which is an n-back task, increases a general cognitive-domain capacity such as fluid intelligence ($G_f$). Thus, the effectiveness of working memory training have been discussed. Although the training’s effectiveness in solving arithmetic word problems needs further investigation, our findings that identified a process in which updating is important would help to develop such training. There is a possibility that the training of the updating function would improve performance on arithmetic word problems, especially in problem solvers who struggle to integrate the information of the problem, as this study indicates that updating is an important process in solving arithmetic word problems.

Our findings also suggest that children with lower updating functions will struggle with integration due to the need to update their problem models. In turn, this suggests a way in which problem solvers with lower updating may be helped. Their performance could be improved when problem model updating is not required, i.e., if problem solvers know how to construct a problem model in advance, they might be able to construct a problem model that includes the relevant information. Indeed, it was reported that problems that use a question as the first sentence are easier (Robinson & Hayes, 1978). In this case, problem solvers could activate their problem schema before reading the problem and then they could integrate the information.
based on the schema. This would reduce the updating load. Therefore, our findings emphasize that when problem solvers have difficulties with updating, it is important to design instruction to reduce the updating that is needed.
References


Appendix A

Examples of arithmetic word problem in Experiment 1.

WP 1. There are five wooden pencils.  
There are 2.6 times as many mechanical pencils as wooden pencils.  
How many mechanical pencils are there?

WP 2. A rectangle is 28 cm long.  
It is 26 cm wide.  
What is its area?

WP 3. A rectangle is 28 cm long.  
A diagonal of this rectangle is 32 cm.  
This rectangle is 15 cm wide.  
What is the rectangle’s area?

WP 4. There are 2.5 l of soy sauce.  
A mirin is 1.4 times as much soy sauce.  
How many liters in the mirin?

WP 5. There are 23 apples.  
There are three times as many oranges as the apples.  
There are eight times as many grapes as apples.  
How many oranges are there?

WP 6. The distance to a destination is 18 km.  
I move at 6 km per hour on foot.  
How many hours do I take to get to the destination?

WP 7. There are four dogs.  
There are five more cats than dogs.  
There are 1.5 times as many birds as dogs.  
How many birds are there?

WP 8. The distance to the destination is 30 km.  
I move at 6 km per hour on foot.  
I move at 15 km per hour by bicycle.  
How long does it take me to get to the destination on foot?

Note. Underlined sentences are extraneous information. These examples were translated into English from Japanese.
Appendix B

Examples of arithmetic word problem in Experiment 2.

| WP 1. | The distance to a destination is 48 km. |
|       | A snake crawls at 7 km/h. |
|       | How many hours does the snake take to get to the destination? |
| WP 2. | A scallop weighs 120 g. |
|       | A cod roe weighs 20 g more than the scallop. |
|       | A soft seaweed weighs 30 g more than the scallop. |
|       | How much does the cod roe weigh? |
| WP 3. | A donkey moves for 4 h. |
|       | The donkey moves at 19 km/h. |
|       | How far does the donkey move? |
| WP 4. | A duck covers 2 km in an hour. |
|       | A gull covers 4 km in an hour. |
|       | The gull moves for 6 h. |
|       | How far does the gull move? |
| WP 5. | There are nine oranges. |
|       | There are three times as many pears as oranges. |
|       | How many pears are there? |
| WP 6. | A lily is 40 cm taller than a dandelion. |
|       | The dandelion is 30 cm. |
|       | How tall is the lily? |
| WP 7. | A distance to a destination was 17 km. |
|       | A rabbit took 8 h to get to the destination. |
|       | A sheep took 6 h to get to the destination. |
|       | How many kilometers an hour did the rabbit cover? |
| WP 8. | There are three times as many scissors as pencils. |
|       | There are six times as many seals as pencils. |
|       | There are four pencils. |
|       | How many seals are there? |

Note. Underlined sentences are extraneous information. A word in necessary information was used as a necessary-information word in the lexical decision task. A word in extraneous information was used as an extraneous-information word. For example, “rabbit” was a necessary-information word and “sheep” was an extraneous-information word in WP 7.
Appendix C

Examples of arithmetic word problem in Experiment 3.

<table>
<thead>
<tr>
<th>Area problems</th>
<th>Length of perimeter problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The vertical line of a square is 10 cm. The horizontal line is 7 cm. The diagonal is 12 cm. What is the area of this square?</td>
<td>1. Side BC of Parallelogram ABCD is 34 cm. A perpendicular from A to side BC is 11 cm. Side AB is 27 cm. How long is the perimeter of this parallelogram?</td>
</tr>
<tr>
<td>2. The vertical line of a square is 14 cm. The horizontal line is 20 cm. The diagonal is 24 cm. What is the area of this square?</td>
<td>2. Side BC of Parallelogram ABCD is 19 cm. A perpendicular from A to side BC is 22 cm. Side AB is 24 cm. How long is the perimeter of this parallelogram?</td>
</tr>
<tr>
<td>3. The vertical line of a square is 20 cm. The horizontal line is 27 cm. The diagonal is 34 cm. What is the area of this square?</td>
<td>3. A right-angled triangle has a vertical line of 12 cm. The length of the horizontal line is 7 cm. The length of the hypotenuse is 14 cm. How long is the perimeter of this triangle?</td>
</tr>
<tr>
<td>4. The vertical line of a square is 36 cm. The horizontal line is 42 cm. The diagonal is 55 cm. What is the area of this square?</td>
<td>4. A right-angled triangle has a vertical line of 18 cm. The length of the horizontal line is 13 cm. The length of the hypotenuse is 22 cm. How long is the perimeter of this triangle?</td>
</tr>
<tr>
<td>5. Side BC of Parallelogram ABCD is 10 cm. A perpendicular from A to side BC is 9 cm. Side AB is 12 cm. What is the area of this parallelogram?</td>
<td>5. Side BC of Parallelogram ABCD is 12 cm. A perpendicular from A to side BC is 14 cm. Side AB is 18 cm. What is the area of this parallelogram?</td>
</tr>
<tr>
<td>6. Side BC of Parallelogram ABCD is 32 cm. A perpendicular from A to side BC is 18 cm. Side AB is 25 cm. What is the area of this parallelogram?</td>
<td>6. Side BC of Parallelogram ABCD is 26 cm. A perpendicular from A to side BC is 15 cm. Side AB is 19 cm. What is the area of this parallelogram?</td>
</tr>
<tr>
<td>7. Side BC of Parallelogram ABCD is 12 cm. A perpendicular from A to side BC is 14 cm. Side AB is 18 cm. What is the area of this parallelogram?</td>
<td>7. Side BC of Parallelogram ABCD is 18 cm. A perpendicular from A to side BC is 24 cm. Side AB is 22 cm. What is the area of this parallelogram?</td>
</tr>
<tr>
<td>8. Side BC of Parallelogram ABCD is 26 cm. A perpendicular from A to side BC is 15 cm. Side AB is 19 cm. What is the area of this parallelogram?</td>
<td>8. Side BC of Parallelogram ABCD is 19 cm. A perpendicular from A to side BC is 22 cm. Side AB is 24 cm. What is the area of this parallelogram?</td>
</tr>
<tr>
<td>9. A right-angled triangle has a vertical line of 10 cm. The length of the horizontal line is 7 cm. The length of the hypotenuse is 12 cm. What is the area of this triangle?</td>
<td>9. Side BC of Parallelogram ABCD is 34 cm. A perpendicular from A to side BC is 11 cm. Side AB is 27 cm. How long is the perimeter of this parallelogram?</td>
</tr>
<tr>
<td>10. A right-angled triangle has a vertical line of 14 cm. The length of the horizontal line is 9 cm. The length of the hypotenuse is 17 cm. What is the area of this triangle?</td>
<td>10. Side BC of Parallelogram ABCD is 19 cm. A perpendicular from A to side BC is 22 cm. Side AB is 24 cm. How long is the perimeter of this parallelogram?</td>
</tr>
<tr>
<td>11. A right-angled triangle has a vertical line of 18 cm. The length of the horizontal line is 24 cm. The length of the hypotenuse is 30 cm. What is the area of this triangle?</td>
<td>11. Side BC of Parallelogram ABCD is 12 cm. A perpendicular from A to side BC is 14 cm. Side AB is 18 cm. What is the area of this parallelogram?</td>
</tr>
<tr>
<td>12. A right-angled triangle has a vertical line of 6 cm. The length of the horizontal line is 4 cm. The length of the hypotenuse is 7 cm. What is the area of this triangle?</td>
<td>12. Side BC of Parallelogram ABCD is 26 cm. A perpendicular from A to side BC is 15 cm. Side AB is 19 cm. What is the area of this parallelogram?</td>
</tr>
<tr>
<td>13. Side BC of ( \triangle ) ABC is 13 cm. A perpendicular from A to side BC is 17 cm. Side AB is 18 cm. What is the area of this triangle?</td>
<td>13. Side BC of ( \triangle ) ABC is 13 cm. A perpendicular from A to side BC is 17 cm. Side AB is 18 cm. What is the area of this triangle?</td>
</tr>
<tr>
<td>14. Side BC of ( \triangle ) ABC is 34 cm. A perpendicular from A to side BC is 21 cm. Side AB is 22 cm. What is the area of this triangle?</td>
<td>14. Side BC of ( \triangle ) ABC is 34 cm. A perpendicular from A to side BC is 21 cm. Side AB is 22 cm. What is the area of this triangle?</td>
</tr>
<tr>
<td>15. Side BC of ( \triangle ) ABC is 7 cm. A perpendicular from A to side BC is 4 cm. Side AB is 6 cm. What is the area of this triangle?</td>
<td>15. Side BC of ( \triangle ) ABC is 7 cm. A perpendicular from A to side BC is 4 cm. Side AB is 6 cm. What is the area of this triangle?</td>
</tr>
<tr>
<td>16. Side BC of ( \triangle ) ABC is 5 cm. A perpendicular from A to side BC is 7 cm. Side AB is 8 cm. What is the area of this triangle?</td>
<td>16. Side BC of ( \triangle ) ABC is 5 cm. A perpendicular from A to side BC is 7 cm. Side AB is 8 cm. What is the area of this triangle?</td>
</tr>
</tbody>
</table>

Note. These are English translations of the problems, which were originally presented in Japanese. In area problems, underlined sentences are extraneous information. Area problems were randomly designated as necessary or extraneous information problems. The order of the second and third sentences in extraneous information problems was also randomized.