Understanding Causal Relations and Learning From Text in Japanese EFL Readers

A Dissertation

Submitted to University of Tsukuba

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Linguistics

Masaya HOSODA

2017

Understanding Causal Relations and Learning From Text in Japanese EFL Readers

A Dissertation

Submitted to University of Tsukuba

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Linguistics

Masaya HOSODA

2017

Abstract

The goal of this dissertation was to explore how Japanese students of English as a foreign language (EFL) understand causal relations and learn from expository text. The ability to learn from expository text in a second language (L2) is important, particularly for university or college students, because this ability can broaden one's expertise beyond the bounds of language and contribute to academic success. In the field of L1 reading research, learning from text is construed as constructing *situation models* of text (Kintsch, 1994). There is general agreement that building coherent situation models largely depends on the understanding of causal relations (*causal understanding*) in text. A body of L1 research has shown that L1 readers implement text comprehension processes during reading (*on-line*) in a manner that is consistent with the causal structure of text (how text information is causally related to other information; Millis et al., 2006; Trabasso & Magliano, 1996). As a result, they build text memory reflecting the text's causal structure after reading (*off-line*; Trabasso & van den Broek, 1985; Varnhagen, 1991).

Unlike L1 research, limited L2 research has investigated L2 readers' understanding of causal relations in expository text and subsequent learning. Therefore, sufficient information is not available about how L2 students' text learning can or should be supported. This dissertation aimed to fill in this gap by addressing the following three general research questions (RQs): (a) What on-line processes and off-line memory are involved in EFL readers' causal understanding of expository text? (b) how does EFL readers' causal understanding contribute to their learning outcomes from expository text? and (c) how do EFL readers' on-line processes and off-line memory reflect the causal structure of expository text?

Study 1 (Experiments 1–3) explored EFL readers' causal understanding and

learning from text. Experiment 1 investigated the relation between EFL readers' causal understanding (assessed by a causal question) and their off-line memory (assessed by a recall test). The results showed that the internal coherence of text memory was more related to causal-question performance than the amount or recalled information, indicating that causal understanding involves memorizing pieces of text information in an interconnected manner.

Subsequently, Experiment 2 explored contributions of causal understanding to learning outcomes from text (assessed by a problem-solving test). It was found that causal understanding contributed to text learning in readers with high L2 reading proficiency, but not in readers with low proficiency. The quantitative and qualitative analyses indicated that low-proficiency readers had trouble both with the understanding of explicit causal relations in text and building accurate situation models of those relations; consequently, they failed to learn the text's causal relations as knowledge.

Experiment 3 investigated whether low-proficiency readers' difficulty with text learning was caused by the applied nature of the problem-solving test (in which readers have to reconstruct and productively transfer text information to a new situation). The results showed that, regardless of L2 reading proficiency, causal-question performance contributed to performance on the problem-solving test that was administered immediately after text reading. This finding supports the idea that low-proficiency readers' difficulty with text learning cannot be simply explained by the nature of the problem-solving test. The qualitative analysis again showed that low-proficiency readers struggled with processes at both textbase and situation-model levels.

Study 2 (Experiments 4–6) was then conducted to gain direct information about on-line reading processes involved in EFL expository comprehension. First, Experiments 4 explored the conditions under which EFL readers make causal bridging inferences that are necessary for building situation models of causal relations in text. The results showed that participants made causal bridging inferences during expository reading on the condition that L2 reading proficiency and the content familiarity of text were both high.

Experiment 5 then employed a self-paced reading method and a recall test to investigate whether and how EFL readers' on-line processes and off-line memory reflected the causal structure of expository text, respectively. Reading times for text statements indicated that participants engaged in on-line processes in a consistent manner with the text's causal structure, when causal relations were made fully explicit. By contrast, the recall results revealed that, regardless of the explicitness of causal relations, participants' off-line memory reflected the causal structure of expository text.

To complement Experiment 5's reading time data, Experiment 6 employed a think-aloud method to explore specific contents of on-line processes during EFL expository reading. Experiment 6 also used the causal question to examine the relation between on-line processes and off-line causal understanding. The think-aloud results first showed that participants causally bridged current and prior text in a manner consistent with the text's causal structure, either when causal relations were made fully explicit or when L2 reading proficiency was high. Regarding the relation between on-line processes and off-line causal understanding, causal-question performance was correlated only with high-proficiency readers' causally bridging current and distant text.

Combining the findings from these six experiments, this dissertation provides a discussion to answer the three general RQs. This study concluded that, in EFL expository comprehension, the relations between (a) on-line processes and causal understanding, (b) causal understanding and learning outcomes from text, and (c) on-line processes and the causal structure of expository text manifest themselves only when the reader and text factors allow for a sufficient amount of attention to be allocated to causal relations across

different parts of the text. This conclusion highlights cognitive characteristics of EFL readers that contrast with those of L1 readers. Finally, I translate the findings into implications for reading instructions that can support EFL readers' expository comprehension and learning from text.

Acknowledgement

This dissertation could not have been completed without support from many people to whom I am indebted. Though it would be difficult to thank all the people who have helped me during my study course at the Graduate School of Humanities and Social Sciences at University of Tsukuba, I would like to extend my special appreciation to the following people.

First, my heartfelt appreciation goes to my chief advisor, Professor Yuji USHIRO, who awakened my interest in foreign language education. He has given me a lot of advice regarding English education and reading research that have made this study what it is, and me who I am today.

I also express my gratitude to Professor Akira KUBOTA for his encouragement and instruction. His comments have improved this study's methodology and interpretation. His lecture on second language acquisition motivated me to explore human language system. My appreciation also goes to Professor Hirosada IWASAKI. Thanks to his teaching, this dissertation's writing quality has largely improved. His academic writing class is always interesting and informative. I also extend my gratitude to Professor Masashi NEGISHI who supervised this dissertation. His comments have informed a better application of this study's findings in light of educational implications; as a result, I managed to clarify contributions of this study to educators in EFL environments, in addition to researchers in the field of reading comprehension.

Many thanks are also due to my colleagues in Professor USHIRO's seminar for their insightful advice and warm support through the course of this study. I also thank Nagomi YAMAKI for her constant support and encouragement. Her presence has given calmness to graduate school days.

Finally, I would like to thank my parents for their irreplaceable support, encouragement, and understanding for attending graduate school.

Masaya HOSODA

Tsukuba, Japan

December, 2017

Table of Contents

List of Tables

List of Figures

Chapter 4

Chapter 5

Chapter 6

List of Appendices

Chapter 1

Introduction

1.1 Educational and Theoretical Backgrounds

1.1.1 Importance of learning from expository text in educational settings

Reading is one of the most powerful tools, through which students acquire much of their knowledge and an understanding of a variety of subject matters (Best, Rowe, Ozuru, & McNamara, 2005; van den Broek, 2010). Particularly relevant to knowledge acquisition is expository text. Some expository texts communicate scientific principles or mechanisms to readers so that they can learn new information as knowledge (Weaver & Kintsch, 1991). The ability to gain new knowledge from expository text written in a second language (L2) is especially important for university or college students because this ability can broaden their expertise beyond the bounds of their native language and contribute to academic progress. Given the increasing demand for English for specific purposes education, it is an important goal of university education to help students learn how to autonomously increase their academic and cultural knowledge from expository text.

However, in English as a foreign language (EFL) education in Japan, there has been insufficient institutional effort to improve students' abilities to learn from text, even in university or college education. Specifically, reading instruction in Japan's English classes often focuses on the interpretation or translation of explicit words or sentences (i.e., learning the text itself), with less attention directed toward training students to learn the new knowledge communicated by the text. To develop instructional practices that can effectively support students' text learning, we need to understand cognitive processes involved in EFL readers' text learning.

1.1.2 Cognitive processes involved in learning from text

In the field of cognitive research on first-language (L1) reading, learning from text is construed as the construction of coherent mental representations of situations described in the text, known as *situation models*) (Kintsch, 1994; McNamara, Kintsch, Songer, & Kintsch, 1996). The construction of situation models involves not only understanding explicit text elements but also inferring relations that connect pieces of text information (Kintsch, 1998). It is well known that causal relations play a central role in the processes of situation-model construction (Magliano, Millis, RSAT Development Team, Levinstein, & Boonthum, 2011; McCrudden, Schraw, & Lehman, 2009; Millis, Magliano, & Todaro, 2006). Causal relations are important for deep comprehension as they provide an essential framework for ordering information in a consistent and logical manner (León & Peñalba, 2002).

In cognitive terms, an understanding of causal relations is defined by an ability to causally explain text (Chi, 2000; McCrudden et al., 2009). L1 research has shown that readers who provide successful causal explanations of text are able to build richer situation models and thus learn more from text (Maglioano et al., 2011; Millis et al., 2006). There is also evidence that L1 readers are more sensitive to the *causal structure* (how text information is causally related to other information) of expository and narrative text during reading (on-line) (Myers, Shinjo, & Duffy, 1987; Radvansky, Tamplin, Armendarez, & Thompson, 2014; Trabasso & Magliano, 1996). They build memory after reading (off-line), reflecting the causal structure of the original text (Trabasso & van den Broek, 1985; Varnhagen, 1990).

In terms of L2 or EFL reading, a number of studies were conducted with narrative text (Muramoto, 2000; Ushiro, Nahatame, Hasegawa, Kimura, Hamada, & Tanaka, 2014; Yoshida, 2003; Zwaan & Brown, 1996). Those narrative studies have shown that

L2 readers keep track of narrative causal structure consisting of characters' goals and actions during reading (Horiba, 1996) and build off-line text memory reflecting the texts' causal structure (Ushiro et al., 2010).

However, compared to narrative text, much less research has been conducted on L2 readers' expository text comprehension. There is no clear understanding of (a) how or to what extent L2 readers understand scientific causal relations in expository text, (b) whether L2 readers' understanding of causal relations (*causal understanding*) contributes to their final learning outcomes, and (c) how EFL readers' on-line reading processes and off-line text memory reflect the causal structure of expository text. Motivated by these questions, this study aimed to reveal cognitive mechanism involved in Japanese EFL readers' understanding of causal relations and learning from expository text. To achieve this goal, I set the following three general research questions (RQs).

General RQ1 What on-line processes and off-line memory are involved in EFL readers' causal understanding of expository text?

General RQ2 How does EFL readers' causal understanding contribute to their learning outcomes from expository text?

General RQ3 How do EFL readers' on-line processes and off-line memory reflect the causal structure of expository text?

These general RQs were addressed by two studies, Study 1 and Study 2. Study 1 investigated EFL readers' causal understanding and learning from text. Study 2 explored on-line processes involved in EFL expository text comprehension. The final goal of this dissertation is to provide empirically grounded findings that inform reading instruction and help promote Japanese EFL readers' acquisition of knowledge through reading.

1.2 Organization of This Dissertation

This dissertation consists of the following six chapters: Introduction (Chapter 1), Literature Review (Chapter 2), Study 1 (Chapter 3), Study 2 (Chapter 4), General Discussion (Chapter 5), and Conclusion (Chapter 6).

In Chapter 2, I first review literature on a cognitive account of reading comprehension. I then review theoretical models and empirical findings regarding the understanding of causal relations, expository text comprehension, and L2 reading. After that, I provide a methodological review of on-line, off-line, and situation-model measures. Last, I summarize the findings and limitations of prior studies and elucidate links between that literature and the present study.

Chapter 3 describes Study 1, consisting of Experiments 1, 2, and 3. Experiment 1 aimed to reveal the relation between text memory and causal understanding in EFL readers. Experiment 2 explored contributions of EFL readers' causal understanding to learning outcomes from text by considering L2 reading proficiency as a variable. Based on Experiment 2's results, Experiment 3 tried to gain a better understanding of the difficulty experienced by low-proficiency EFL readers in text learning.

Chapter 4 describes Study 2, consisting of Experiments 4, 5, and 6. Experiment 4 sought to clarify the conditions under which EFL readers make causal bridging inferences during expository reading that are necessary for building causal relations. Experiments 5 and 6 explored the relation between EFL readers' on-line processes and the causal structure of the expository text by using a self-paced reading and a think-aloud method, respectively. Experiment 6 also investigated how EFL readers' on-line processes relate to causal understanding after reading.

Chapter 5 combines findings from the six experiments in the general discussion of this research. I discuss the present finding in terms of (a) on-line processes and off-line memory involved in EFL readers' causal understanding, (b) the relation between EFL readers' causal understanding and learning from text, and (c) how EFL readers' off-line memory and on-line processes reflect the causal structure of the expository text.

Chapter 6 summarizes the findings from the experiments with reference to the three General RQs. I then discuss limitations of the study and provide suggestions for future research. Finally, I suggest the implications of this study's findings for reading instruction that assists students' learning from text.

Chapter 2

Literature Review

2.1 A Cognitive Account of Reading Comprehension

2.1.1 Situation model framework and learning from text

In the first part of this literature review, and to mark the starting point of this dissertation, I will provide a cognitive account of reading. Accumulating cognitive reading research over the past 40 years has indicated that what remains in readers' memory after reading is not limited to what is explicit in the text (e.g., Kintsch, 1998; McNamara & Kintsch, 1996; McNamara et al., 1996). One important idea of cognitive accounts of reading comprehension is that mental representations of text constructed through reading are not properties of the text itself, but cognitive structures built by the reader on the basis of the text (Kintsch, 1994, 1998; Kintsch & van Dijk, 1976; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). While advancing through text, readers activate their prior knowledge based on information explicitly provided in the text (e.g., Kintsch, 1998). This process builds connections between the information being read and previous parts of the text, as well as between current information and readers' prior knowledge (Graesser & Bertus, 1998; Graesser, Singer, & Trabasso, 1994; Kintsch, 1998; Singer, Halldorson, Lear, & Andrusiak, 1992; Singer, Harkness, & Stewart, 1997).

One widely used metaphor in this context is that text comprehension amounts to the construction of mental representations of what the text is about (e.g., Kintsch, 1994, 1998; Kintsch & van Dijk, 1976; van Dijk & Kintsch, 1983). It has been agreed that readers construct at least three mental representations in the course of text comprehension: *surface memory*, *textbases*, and *situation models* (e.g., Kintsch, 1994; Kintsch, 1998; McNamara & Kintsch, 1996; McNamara et al., 1996).

Surface memory is verbatim memory of individual explicit words or phrases. At this level, mental representations do not have true-false values and are susceptible to time decay (Kintsch, Welsch, Schmalhofer, & Zimmy, 1990). Hence, surface memory is of little relevance to the discourse processes considered here.

Textbase is a representation of meanings of text, comprised of textual propositions and their interconnections. Textbase is relevant to discourse processes because it include the understanding of relations connecting information in text (e.g., Kintsch, 1998; McNamara & Kintsch, 1996). Building textbase involves readers capturing meanings of individual textual elements and integrating them into a meaningful proposition.

Finally, situation models are referential representations of situations described in text, amounting to an amalgamation of the explicit text and readers' prior knowledge in long-term memory (Kintsch, 1994, 1998; McNamara et al., 1996; van den Broek, Lorch, Linderholm, & Gustafson, 2001). Figure 2.1 shows an example of a situation model of a text about heart disease.

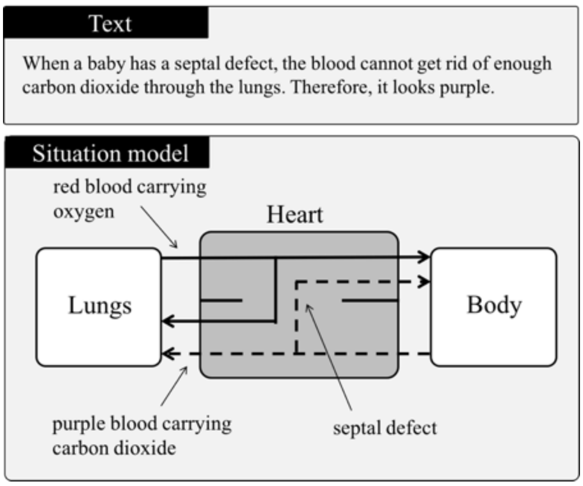

Figure 2.1. Textbase and a situation model for a two-sentence text on heart disease (adapted from Kintsch, 1994, p. 295).

As stated in the introduction, learning from text is construed as the construction of situation models (Goldman, 1997; Kintsch, 1994; McNamara et al., 1996). At the situation-model level, readers go beyond the explicit text to understand underlying relations, principles, or mechanisms that explain how and why individual text events lead to and result from other events (Kintsch, 1998). Thus, when situation models are successfully constructed, readers are assumed to use what they have learned from the text as new knowledge for applied purposes (e.g., providing explanations or solving problems).

One important aspect of the nature of situation models lies in their flexibility; the lack of restriction from explicit text enables situation models to take various forms according to factors brought into reading (e.g., McNamara & Kintsch, 1996). Critical factors determining the richness of situation models include readers' *standards of coherence* (e.g., van den Broek et al., 2001; van den Broek, Virtue, Everson, Tzeng, & Sung, 2002). Standards of coherence refer to "the degree of comprehension that a reader attempts to achieve during the reading of a text" (Linderholm, Virtue, Tzeng, & van den Broek, 2004, p. 168). In terms of standards of coherence, situation models are enriched only when readers base their "standards" on the understanding of situations conveyed by a text (not the text itself)—readers are not satisfied with their comprehension until they successfully capture the "world" or "messages" communicated by the author. In contrast, with standards of coherence placed on the interpretation of explicit information, readers are likely to settle at the point where they understand meanings of explicit words or sentences. Therefore, the building of situation models inherently depends on readers' standards of coherence.

2.1.2 Role of inference generation in building coherence

As described above, cognitive accounts of reading comprehension posit that the successful construction of situation models allows readers to acquire text information as new knowledge (e.g., Kintsch, 1994; McNamara et al., 1996). However, it is widely agreed that what is explicitly stated in text is usually insufficient to construct coherent situation models (Britton, 1994; Kintsch, 1998; McNamara et al., 1996). When explicit information is insufficient for situational understanding, readers must infer additional propositions or links that are necessary to fill conceptual gaps in explicit information (Graesser, Bertus, & Magliano, 1995; Graesser et al., 1994; Kintsch, 1998; McNamara & Kintsch, 1996; McNamara et al., 1996; Singer et al., 1992; Singer et al., 1997). There is a consensus that this *inference generation* is the most central cognitive process in the construction of situation models (e.g., Graesser et al., 1994; Horiba, 2000; Kintsch, 1994; McNamara et al., 1996; Singer et al., 1992; Singer et al., 1997; van den Broek et al., 2002). Considerable evidence has shown that inferences play a critical role in establishing links between pieces of text (Myers et al., 1987; Wiley & Myers, 2003) and between text information and readers' prior knowledge (Singer et al., 1992; Singer et al., 1997; van den Broek et al., 2002), in addition to maintaining coherence of discourse (e.g., Kintsch, 1994; O'Reilly & McNamara, 2007; McNamara et al., 1996). Due to this significance of inferences, the ability of appropriately making inferences is known to be a key characteristic of skilled readers (Carlson et al., 2014; Horiba, 2000; Li & Kirby, 2014; Zwaan & Brown, 1996).

To date, various discourse processing models have been advocated to explain categories of inference that can be generated during (natural) reading, such as the *minimalist hypothesis* (e.g., McKoon & Ratcliff, 1992), the *constructionist theory* (e.g., Graesser et al., 1994), the *memory-based view* (O'Brien, Rizzella, Albrecht, & Halleran, 1998), and the *landscape model* (van den Broek et al., 2002). A detailed review of these discourse models is beyond the scope of this study. For present purposes, it should be noted that these models generally agree that inferences necessary for maintaining coherence between adjacent sentences (known as *local coherence*) are generated during reading (Graesser et al., 1994; Singer et al., 1992; Singer et al., 1997).

Bridging inferences are a typical type of such inferences. Bridging inferences are generated to link text information with implicit ideas (e.g., Noordman et al., 1992; Singer et al., 1992; Singer et al., 1997). Consider the sequence, "A burning cigarette was carelessly discarded. The fire destroyed many acres of virgin forest." Coherence between the two sentences cannot be established unless readers infer an additional proposition that connects the events described in the sentences, such as "the burning cigarette made a fire." Therefore, the essence of bridging inferences is that they serve to maintain coherence between parts of text when explicit text information is insufficient to do so (Kintsch, 1998; Singer et al., 1992). If necessary bridging inferences are not made, incoherence emerges, and situational understanding would be impossible. These circumstances point to the fact that coherence is a result of bridging inferences. Hence, the building of situation models depends on inference generation by readers, in addition to information explicitly present in text.

To summarize the above cognitive account of reading, reading comprehension can be defined as a cognitive activity to build mental representations of situations described in text (e.g., Kintsch, 1998). This process of situation-model construction is, to a large extent, dependent on readers' inference generation (e.g., Graesser et al., 1994). To further explore the scope of this study, the following sections describe the (a) roles of causal relations in text comprehension, (b) characteristics of expository text comprehension, and (c) nature of L2 reading.

2.1.3 Role of causal relations in text comprehension

In the course of text comprehension, readers understand not only individual pieces of text information but also relations connecting text information to form a coherent whole (e.g., McNamara et al., 1996; Millis et al., 2006; Singer et al., 1992; Singer et al., 1997). Discourse processing researchers have proposed that information in situation models is interconnected by means of at least five dimensions: causal, temporal, spatial, logical, and intentional relations (e.g., Zwaan & Radvansky, 1998). Among these, none has received more scrutiny than causal relations in cognitive research on L1 and L2 reading (Langston, Trabasso, & Magliano, 1999; McCrudden et al., 2009; Millis et al., 2006; Singer et al., 1992; Singer et al., 1997; Trabasso, van den Broek, & Suh, 1989; Ushiro et al., 2010; Ushiro et al., 2015; Varnhagen, 1991).

There is plenty of evidence that situation-model construction largely depends on readers' ability to detect and understand causal relations in text (McCrudden et al., 2009; Millis et al., 2006; Mulder, 2008). Specifically, researchers have demonstrated that (a) causally related information is processed faster and encoded into memory more strongly than information conveyed by other types of relations (Mulder, 2008; Singer et al., 1992), (b) elaborating text by signaling causal relations enhances comprehension (Linderholm, Everson, van den Broek, Mischinski, Crittenden, & Samuels, 2000), (c) skilled L1 readers base their standards of coherence on causality among multiple sentences (Wittwer & Ihme, 2014), and (d) the ability to identify causal relations between sentences during reading discriminates between skilled and less skilled L1 readers (Gilliam, Magliano, Millis, Levinstein, & Boonthum, 2007; Magliano et al., 2011).

In particular, there is agreement among past L1 reading studies that how well readers can causally explain text is directly linked to their comprehension at a deeper level (Chi, 2000; Millis et al., 2006; León & Peñalba, 2002; Trabasso, Suh, Payton, &

Jain, 1995). Causal explanation involves the integration of bits of text information into one coherent idea (Coté, Goldman, & Saul, 1998; Millis et al., 2006). Therefore, to causally explain a text, readers must not only memorize relevant information from the text but also have an integrative understanding of the series of causal relations between pieces of that information (Coté et al., 1998; Magliano & Millis, 2003; Magliano et al., 2011). Therefore, causal explanation can be a good indicator of readers' understanding of a text's causal relations. Stated differently, causal understanding of text can be defined by the ability to causally explain it.

The view that reading is explanation based is supported by L1 studies demonstrating that comprehension is guided by readers' *causal reasoning* (Graesser et al., 1994; Magliano et al., 1999; Trabasso & Suh, 1993). Causal reasoning is an activity where readers try to causally explain current information based on prior texts as well as their own prior knowledge. One study frequently cited in this discussion is that of Myers et al. (1987). They used sentence pairs like those in Table 2.1 as experimental material. These sentence pairs were manipulated according to the strength of causal relations between events in sentences.

Table 2.1

Sentence Pairs With Causal Relations of Different Strengths

1.	Joey's brother punched him again and again. The next day his body was covered with bruises.
2.	Racing down the hill, Joey fell of his bike. The next day his body was covered with bruises.
3.	Joey's crazy mother became furiously angry with him. The next day his body was covered with bruises.
4.	Joey went to a neighbor's house to play. The next day his body was covered with bruises.

In the first sentence pair, it can be readily interpreted that the event described in the first sentence (Joey's brother punching him) is the cause of the circumstances described in the second sentence (Joey having bruises). On the other hand, the strength of causal relations of the events in the third sentence pair is moderate; the causal interpretation in this case requires readers to infer an additional proposition (e.g., "Joey's mother punched him") that explains how the events in the sentences can be connected. Finally, for the events in the fourth sentence pair, causal interpretation seems implausible, meaning that they are causally unrelated.

Myers et al.'s (1987) study revealed two critical findings. First, reading times for second sentences became shorter as a function of the strength of causal relations. This observation indicates that readers understood second sentences based on their causal relations to first sentences, and consequently, on-line processes were facilitated by the strength of those relations. Second, readers recalled events that were moderately causally related (i.e., Sentence Pair 3 in Table 2.1) better than either events that are strongly related (i.e., Sentence Pair 1 in Table 2.1) or events that are unrelated (Sentence Pair 4 in Table 2.1). The enhanced recall of moderately related pairs was assumed to be a result of inference generation associated with the understanding of those moderate pairs (Golding, Millis, Hauselt, & Sego, 1995). Together, these findings lead to the conclusion that causal relations play a primary role in text comprehension processes.

Based on this importance of causal relations, L1 researchers later developed the *causal network model*, which systematically describes how text's causal relations affect readers' on-line processes and ensuing off-line memory of the text (Langston et al., 1999; Suh & Trabasso, 1993; Trabasso & Sperry, 1985; Trabasso & van den Broek, 1985; Trabasso et al., 1989). Along with reflecting the findings from Myers et al.'s (1987) study, the causal network model is based on the assumption that comprehension is guided by causal reasoning. One important contribution of the causal network model is its principled discourse analysis system for identifying causal relations between statements in text. To be more precise, the causal network model identifies causal relations in text based on *logical criteria of necessity* and *weak sufficiency*. Event A is necessary for Event B if Event B does not occur in the described text situation without Event A. Furthermore, Event A is sufficient for Event B if Event B is likely to follow Event A.

Consider Magliano and Millis's (2003, p. 256) statements, for example. Figure 2.2 illustrates the causal relations between these sentences, with each relation expressed with the arrow. Causal network analysis judges Sentences 2, 3, and 4 to be directly related to Sentence 5. In other words, the events in all three sentences are necessary for the event in Sentence 5 to occur; had the rich lord not liked to collect carvings of animals, had a mouse carving, or not called the skilled carvers, then he would not have asked the carvers to make a carving of a mouse. Additionally, analysis judges that Sentence 2, 3, and 4 events are jointly sufficient for the Sentence 5 event because it is highly likely that the rich lord would ask for a carving of a mouse if all events stated in Sentences 2, 3, and 4 were to occur.

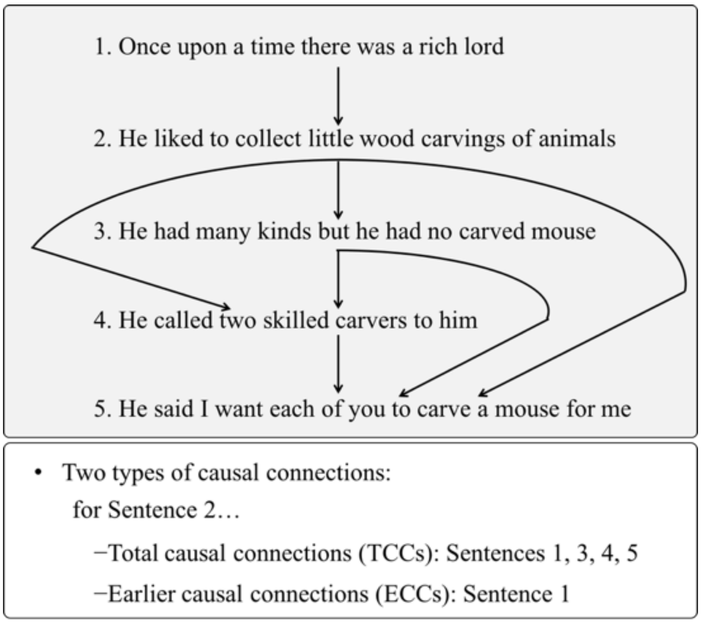

Figure 2.2. A causal structure of the five sentences adapted from Magliano and Millis (2003, p. 256).

Therefore, performing causal network analysis yields a network of causal relations where some text statements have more causal connections than others. This networked structure is called the *causal structure* (e.g., Suh & Trabasso, 1993; Trabasso & Suh, 1993). Information from the causal structure can then be used to predict on-line processes as well as off-line text memory. There are three aspects that are particularly relevant to the present study.

First, the causal network model predicts that statements with more causal connections to other statements in a text are better recalled than statements with fewer causal connections (Trabasso & van den Broek, 1985; Ushiro et al., 2010; Ushiro et al., 2015). The number of causal connections that any statement shares with the entire text is referred to as *total causal connections* (TCCs) in this study. For example, Sentence 2 is expected to be the best recalled of the five sentences in Figure 2.2 because it has the highest number of TCCs (i.e., four TCCs). In essence, information with many TCCs is presumably associated with increased retrieval access, and is therefore better remembered later (e.g., Radvansky et al., 2014). Enhanced recall for information with many TCCs in comparison to information with fewer TCCs is interpreted as evidence that readers' off-line memory reflects the causal structure of the text in question (Horiba, 1996; Trabasso & van den Broek, 1985; Ushiro et al., 2010). Indeed, there has been strong empirical support for this model in both L1 and L2 studies (Horiba, 1996; Horiba, van den Broek, & Fletcher, 1993; Radvansky et al., 2014).

Second, the causal network model predicts that readers should make causal links between current statements and prior statements when there are causal connections between them (Magliano et al., 1999; Millis et al., 2006; Suh & Trabasso, 1993). The causal connections that any information shares with earlier parts of a text are called e*arlier causal connections* (ECCs). In Figure 2.2's example, readers should causally bridge Sentence 5 to Sentence 3 through such reasoning as "the rich lord requested mouse carvings (Sentence 5) because he did not have a carved mouse (Sentence 3)." This type of processing is called *causal bridging* because it causally connects information units explicitly stated in text to each other. When readers engage in causal bridging more frequently for information with many ECCs than for information with fewer ECCs, their on-line processes are interpreted to reflect the causal structure of the text they read. This occurrence is exactly what past studies have found in L1 readers of narrative text (Horiba, 1996; Magliano et al., 1999; Suh & Trabasso, 1993; Trabasso & Magliano, 1996), though there is no evidence of this processing in L2 readers.

Third, it is assumed that readers take less time to read statements with many ECCs than statements with fewer ECCs (Magliano et al., 1999; Radvansky et al., 2014), because information that can be more easily integrated with a larger part of text that has already been read should be easier to process, leading to a decrease in reading time. As opposed to the first and second prediction, this third prediction regarding processing times has not been addressed appreciably in past studies. Although a decrease in reading times for increased ECCs was reported in several L1 studies using narrative texts (e.g., Magliano et al., 1999), some researchers claim that increased ECCs do not always shorten reading times. For example, there are situations where readers cannot readily detect causal relations between current information and prior text because those relations remain implicit in the text (Trabasso & Suh, 1993). In that case, reading time may instead be longer for information with more ECCs, as its understanding should involve readers' inference generation about causal relations.

The causal network model accordingly provides an account of the relation between text's causal structure and readers' behavioral measures. However, it is important to note that the causal network model does not provide a direct explanation of how text's causal relations are integrated with readers' prior knowledge to form situation models. Rather, this model specifically focuses on processes to bridge pieces of explicit text information to each other (in particular, causal bridging).

For this point, Singer and his colleagues have provided a theoretical framework called the *validation model*, which posits that the building of situation models of causal relations involves making bridging inferences about those causal relations (Halldorson & Singer, 2002; Singer & Halldorson, 1996; Singer et al., 1992). Specifically, they suggested that before causal relations are accepted by readers to form situation models, readers should infer linking propositions that explain how a given cause leads to its effect.

To illustrate, consider the sequence, “Dorothy poured the bucket of water on the bonfire. The fire went out.” Upon encountering this sequence, readers should readily detect a possible causal relation between the two sentences: pouring the water *caused* the fire to go out. In addition, situational understanding of this sequence is assumed to involve readers’ inferring a proposition, “water extinguishes fire,” which explains why the possible cause (water was poured on the bonfire) brings about the effect (the fire went out) described in the sentences. These inferences are referred to as *causal bridging inferences* because they causally bridge pieces of information with an implicit linking proposition (e.g., Noordman et al., 1992; Singer et al., 1997). To avoid any confusion, it should be noted that I use the term “causal bridging inferences” to describe inferential processes to derive additional (not explicitly stated) propositions about a given causal relation, whereas the term “causal bridging” (introduced in the review of the causal network theory) is used to describe processes to link explicit text propositions to each other.

As shown in the above example, if causal bridging inferences are not generated, the understanding of a causal relation remains at the level of explicit information (i.e., textbase). Therefore, causal bridging inferences are necessary for building situation models of text’s causal relations (Best et al., 2005; Noordman & Vonk, 1998; Singer et al., 1997). In a series of experiments with narrative text, Singer and his colleagues found empirical support for L1 readers making causal bridging inferences during reading (Halldorson & Singer, 2002; Singer & Halldorson, 1996; Singer et al., 1992). Singer et al. (1997) attempted to extend this finding to expository text—the details of their findings are described later.

In sum, the theoretical discourse models and empirical research findings together recognize the notion that understanding causal relations plays an important role in situation-model construction (Chi, 2000; Millis et al., 2006; Singer et al., 1997; Trabasso et al., 1989). From this perspective, one may suspect that causal understanding contributes to learning from text in EFL readers. However, little research has directly explored this issue. To establish the background of L2 expository reading, the following two sections review previous findings regarding expository text and L2 reading comprehension.

2.1.4 Expository text comprehension

2.1.4.1 Nature of the expository text and key cognitive processes involved in expository comprehension

The main goal of expository text is to communicate information to readers so that they can acquire new knowledge from the text (Best et al., 2005; Weaver & Kintsch, 1991). Although past research has regarded history or social study passages as expository text (Linderholm et al., 2000; Voss & Silifies, 1996), this study uses the term "expository text" to refer to scientific passages describing a series of events (McCrudden et al., 2009; Ushiro et al., 2015). It should also be noted that when the single word "text" is used in this dissertation, it refers to expository text unless otherwise noted.

Before exploring the nature of expository text comprehension, it is helpful to briefly consider narrative text. Narrative text is written mainly to entertain readers. Accordingly, narrative text conveys stories that are rooted in typical readers' everyday experiences, such as characters' goals, actions, and emotions (Mandler & Johnson, 1977; Trabasso & van den Broek, 1985). Of primary relevance to this study, causal relations in narrative text are associated with characters' goals and actions (Graesser & Clark, 1985; Trabasso et al., 1989). It is widely agreed that readers have a rich amount of world knowledge, or schemata, about goal-action relations because taking actions to achieve goals is quite common in everyday life (Langston et al., 1999). Therefore, understanding causal relations in narrative text is, to a large extent, guided by the passive or automatic activation of readers' schematic knowledge about stories (Gerrig, 2005; McKoon & Ratcliff, 1992; Zwaan & Radvansky, 1998).

By contrast, because expository text is written to provide new information on a certain topic or domain, causal relations in expository text often consist of sequences of physical or scientific events (Noordman et al., 1992; Singer et al., 1997). These specifically include mechanisms about how scientific phenomena occur, how physical components in specific systems work, and how such systems are used for specific purposes (Goldman, 1997; Weaver & Kintsch, 1991). Because these relations are often abstract or technical, especially when compared to goal-action relations in narratives, a robust knowledge base is less available for understanding expository text (Best et al., 2005; Goldman, 1997). Therefore, expository text comprehension is less guided by knowledge-based passive processes (Best, Floyd, McNamara, 2008; Singer et al., 1997). Indeed, prior expository comprehension studies have shown that readers of expository text must take conscious control, for textbase-level processes to make meaning from individual words or sentences and to link individual text elements to form propositions, as well as situation-model-level processes (Coté et al., 1998; Horiba, 2000). Because situation-model construction is generally established on the appropriate building of a textbase (e.g., Kintsch, 1998), readers of expository text must tackle these textbase processes before they can build coherent situation models and learn from text.

One important process of building robust textbase for expository text is causal bridging (e.g., León & Peñalba, 2002; Magliano et al., 2011; Millis et al., 2006; van den Broek et al., 2002). The importance of causal bridging in expository comprehension can be well understood through findings by Magliano and colleagues (e.g., Gilliam et al., 2007; Magliano et al., 2011; Millis et al., 2006). These researchers performed causal network analysis on expository texts, thereby identifying where L1 readers are theoretically assumed to engage in causal bridging during reading (i.e., statements with many ECCs). In one study, they found that the frequency of causal bridging for a theoretically identified part of text was significantly correlated with scores on a standardized reading test (Magliano et al., 2011). This finding supports the notion that causal bridging is a characteristic of skilled readers (e.g., Suh & Trabasso, 1993). More important, they found that L1 readers' causal bridging of information from a current sentence to information from a distant part of a text (termed *distal bridging*) predicted how well readers causally explained content of texts, as assessed by a question probing causal relations (the *causal question*, Millis et al., 2006). This observation proposes that causal bridging implemented during L1 expository reading is linked to causal understanding of expository text.

2.1.4.2 Causal inferences during expository reading

As mentioned in Section 2.1.2, readers must usually make inferences in order to bridge pieces of information in text. However, the generation of bridging inferences is known to be more difficult for expository text than for narrative text (e.g., Best et al., 2005; Shimizu, 2015). This difficulty is mainly because of expository text's less familiar content (e.g., Noordman & Vonk, 1992; Noordman et al., 1992; Singer et al., 1997), which tends to force readers to concentrate on analyzing individual words/sentences. It has been demonstrated that even L1 readers do not always make causal bridging inferences necessary for maintaining local coherence of text during expository reading. Noordman et al. (1992) conducted a pioneering study on the generation of causal bridging inferences from expository text. They used technical scientific passages and examined whether L1 university students make causal bridging inferences during reading. Specifically, their experimental passages included *because* target sentences such as "Chlorine compounds make good propellants because they react with almost no other substances". In this sentence, the causal connective *because* signals the presence of a causal relation; the subordinate clause (chlorine compounds react with almost no other substances) is a cause of the main clause (chlorine compounds make good propellants). What is left implicit is the proposition "propellants must not react with the materials in a spray can," which explains why the cause leads to the effect. The fully situational understanding of this causal relation requires readers to derive this mediating proposition through the generation of causal bridging inferences. Unexpectedly, the results from a series of experiments showed that L1 readers did not make causal bridging inferences during expository reading unless they were required to do so in a post-reading task.

On the other hand, Singer et al. (1997) found supportive evidence for L1 readers' generation of causal bridging inferences with relatively familiar expository text, but on the condition that causal relations were signaled by the causal connective *because*. Table 2.2 summarizes previous L1 research findings regarding the generation of causal bridging inferences from expository text.

Table 2.2

Summary of Findings From Past L1 Studies on Causal Bridging Inferences During Expository Text Reading

Literature	Factor(s)	Do readers make causal bridging inferences during reading?
Noordman and Vonk (1992)	The amount of prior knowledge about text's subject matter	YES for readers with high background knowledge NO for readers with low background knowledge
Noordman, Vonk, and Kempff (1992)	The presence of a reading goal to detect inconsistencies in texts	NO except when readers were given the reading goal to detect inconsistencies in texts was given
Singer, Stewart, and Harkness (1997)	The familiarity of text content	YES for familiar texts NO for unfamiliar texts
Singer and Gagnon (1999)		YES
Singer and O'Connell (2003)	The presence of the causal connective *because*	YES when causal relations were signaled by the causal connectives NO when causal relations were not signaled
Wiley and Myers (2003)	Availability (presence in texts) and accessibility (distance from target sentences) of information necessary for inference generation	YES only when all the necessary information is both available and accessible.

These L1 studies converge to suggest that L1 readers make causal bridging inferences during expository reading when two conditions are met: (a) the text's content is not too technical to readers, and (b) causal relations between critical statements are signaled by causal connectives.

Accordingly, even L1 readers do not always generate inferences that are necessary for building situation models of causal relations—what about L2 reading? Is it difficult or even impossible for L2 readers to make inferences and learn from expository text? To build the theoretical and empirical backgrounds to answer these questions, the following three sections review the cognitive nature of L2 reading comprehension.

2.2 Reading in a Second Language

2.2.1 Differences from L1 reading

Cognitive processes involved in text comprehension are generally common between L1 and L2 readers (Grabe, 2009). When L2 readers comprehend text, they first engage in lower-level linguistic processes, such as recognizing strings of letters as a word, capturing meaning of individual words and sentences, and integrating individual text elements into one proposition (Koda, 2005). In addition, L2 readers (especially proficient ones) may engage in higher-level conceptual processes to build a situational understanding of text, such as monitoring the degree of their comprehension, making inferences from prior knowledge, and continuously updating evolving mental representations according to incoming information (Hosoda, 2014; Horiba, 2000, 2013).

However, an important difference between L1 and L2 readers lies in their efficiency of lower-level processes (Grabe, 2009; Horiba, 2000; Zwaan & Brown, 1996). Usually, lower-level processes in L2 readers are less automatized than in L1 readers, thus drawing many of their cognitive resources. This situation is theoretically explained within the framework of the *cognitive capacity theory* (Just & Carpenter, 1992). The cognitive capacity theory posits that the amount of cognitive resources available at any time point of reading is limited; various levels of cognitive processes therefore compete for the limited resources during text comprehension. One critical assumption is that if lower-level processes demand a large amount of cognitive resources, those processes are prioritized over higher-level processes in the allocation of readers' cognitive resources. These circumstances have been widely demonstrated by past L2 reading studies (Horiba, 1996; Horiba, 2000; Zwaan & Brown, 1996). Specifically, these L2 studies reported that the high cognitive demands of lower-level processes contribute to two major characteristics of L2 reading.

First, L2 readers engage in higher-level conceptual processes to a lesser extent than do L1 readers. For example, two think-aloud experiments by Horiba (2000 [Experiment 1], 2013) showed that L2 readers paid most attention to interpreting explicit meaning of currently read words or sentences, engaging in higher-level processes (e.g., inference generation and association with prior knowledge) less than L1 readers. Along the same line, Zwaan and Brown (1996) showed that the frequency of inference generation largely dropped when participants read in L2 rather than L1.

Second, on-line processes in L2 readers do not necessarily directly contribute to off-line memory or comprehension. The underlying rationale is that lower-level processes are necessary but not sufficient for deep comprehension (e.g., Kintsch, 1998). L2 readers' excessive focus on lower-level processes often overrides contributions of integrative relational processes to ensuing memory or comprehension. For example, Horiba (2013 [Experiment 2]) reported that any type of think-aloud verbal protocols did not correlate with performance on a recall task, thus indicating that the relations between on-line processes and off-line memory in L2 reading are not straightforward, and more complex than in L1 reading.

2.2.2 Inference generation in L2 reading

Collectively, the L2 findings reviewed in the previous section indicate that L2 readers are distinctive from L1 readers with regard to their limited efficiency of lower-level processes. At the same time, it is unlikely that all of the activity that L2 readers undertake during reading is done within the confines of lower-level processes. As stated at the beginning of this section, building situation models in L2 reading involves readers' engagement in higher-level processes, including inference generation, just as it does for L1 readers.

Several recent L2 studies were conducted to explore higher-level processes (Hosoda, 2014; Horiba, 1996, 2000, 2013; Kato, 2014; Shimizu, 2015; Yoshida, 2003). Those studies often focused on inference generation (e.g., Hosoda, 2014; Muramoto, 2000; Shimizu, 2015; Yoshida, 2003), mainly due to its importance in the construction of situation models (e.g., Kintsch, 1998). An important finding is that L2 readers generate inferences during reading when their L2 reading proficiency is high enough (although to a limited extent when compared to L1 readers; Horiba, 1996; Hosoda, 2014; Muramoto, 2000). The basic idea underlying this finding is that readers with high L2 reading proficiency are usually able to complete lower-level processes with just a few cognitive resources, thus leaving resources available for inference generation.

For example, in her think-aloud experiment, Yoshida (2003) reported that EFL readers with high L2 reading proficiency made more elaborative inferences while reading narrative texts than low-proficiency readers. In addition, Horiba (1996) showed that advanced L2 readers of narrative texts generated more elaborative and bridging inferences than did intermediate L2 readers, who were found to focus on information they were currently reading. Furthermore, Muramoto (2000) demonstrated that, after reading narrative texts, proficient EFL readers falsely recognized unstated but inferable information more frequently than did less proficient readers. He interpreted this result as showing that proficient readers inferred implicit information during reading, which was then encoded into their memory representations.

Taken together, these studies point to the fact that how well L2 readers make inferences during reading is dependent on L2 reading proficiency. If they have sufficiently high proficiency that allows for the efficient operation of lower-level processes, L2 readers can potentially engage in inferential processes and construct situation models. Based on this view, it is likely that proficient L2 readers can make causal bridging inferences and build situation models of those causal relations. Conversely, due to the difficulty of inferential processing, less proficient L2 readers may struggle to construct situation models of causal relations, which could prevent them from acquiring causal relations as knowledge.

At the same time, it must be noted that only a few studies (Hosoda, 2014; Shimizu, 2015) have examined L2 readers' inference generation using expository text. It thus remains unclear when or how L2 readers make causal bridging inferences, which are necessary for building situation models of causal relations, while reading expository text. Based on previous L2 narrative studies and L1 expository studies, potential factors affecting L2 inference generation from expository text include L2 reading proficiency (e.g., Muramoto, 2000) and the familiarity of text content (e.g., Singer et al., 1997). To address L2 readers' causal understanding and learning from text, research must consider these factors and clarify under what conditions L2 readers make causal bridging inferences during expository reading.

2.2.3 Understanding causal relations in L2 reading

One major goal of this dissertation is to explore how EFL readers understand causal relations in expository text. This section reviews existing findings regarding L2 readers' on-line processes and off-line memory for causal relations in text. First, it must be noted that most past studies on the understanding of causal relations in L2 reading were conducted with narrative text (Horiba, 1996; Ushiro et al., 2010; Ushiro et al., 2015); this line of L2 research using expository text is quite limited (Hosoda, 2014; Ushiro et al., 2015). Therefore, I first review findings from past L2 narrative studies to provide an initial overview of L2 readers' causal understanding. Then, I shed light on a few studies that investigated L2 readers' understanding of causal relations in expository text.

Several empirical studies were conducted in which researchers applied the causal network model to L2 narrative text comprehension (e.g., Horiba, 1996; Horiba, van den Broek, & Fletcher, 1993; Ushiro et al., 2010). Table 2.3 summarizes findings from past L2 research using the causal network model. These findings demonstrate that L2 readers' on-line processes and off-line memory usually reflect narrative causal structure consisting of characters' goals and actions, although to a limited extent compared to L1 readers' processes.

Table 2.3

Summary of Findings From Pas L2 Studies Using the Causal Network Model

	Text type	Measure(s)	Key results
Horiba et al. (1993)	Narrative	Recall	Recall rates increased as a function of the number of TCCs.
Horiba (1996)	Narrative	Recall Think-aloud	Recall rates increased as a function of the number of TCCs only for L2-advanced readers and L1 readers. Inference generation did not increase as a function of the number of TCCs in L2 readers. L2-advanced readers made inferences at a later part of the texts.
Ushiro et al. (2010)	Narrative	Recall Reading times	Recall rates increased for statements with two and three or more TCCs, compared to statements with no and one TCC. Reading times were shorter for statements with two and three or more TCCs than statements with one and no TCC.
Ushiro et al. (2015)	Expository	Recall	Recall rates were highest for statements with highest numbers of TCCs, followed by middle numbers of TCCs, and lowest numbers of TCCs.

Note. TCCs = total causal connections (i.e., the number of causal connections any information has to the other information in the text).

Regarding characters' goal-action relations, researchers also found that EFL readers inferred characters' goals that were not explicitly stated in text and encoded them into memory representations (Ushiro et al., 2014), meaning that EFL readers have schematic knowledge about goal-action relations that allows for the inference and encoding of implicit goal information. Thus, it can be concluded that L2 readers are generally sensitive to narrative causal relations that consist of characters' goals and actions.

However, it must be noted that there is a critical limitation in past L2 studies using the causal network model. Although these studies targeted on-line processes, they looked at the effect of information's causal connections to an entire text (i.e., TCCs). Consider Figure 2.3. In this case, past L2 studies would have judged that current information (3) has four causal connections (1 + 2 + 4 + 5), and used this as what would affect the on-line processing of 3. However, as Figure 2.3 shows, causal connections that are available for the processing of current information are 1 and 2, located earlier in the text (i.e., ECCs; Radvansky et al., 2014). Stated differently, 3 and 4, located later in the text, are not available for current processing because they are yet to be encountered. Thus, to examine on-line processes, research should use ECCs, not TCCs.

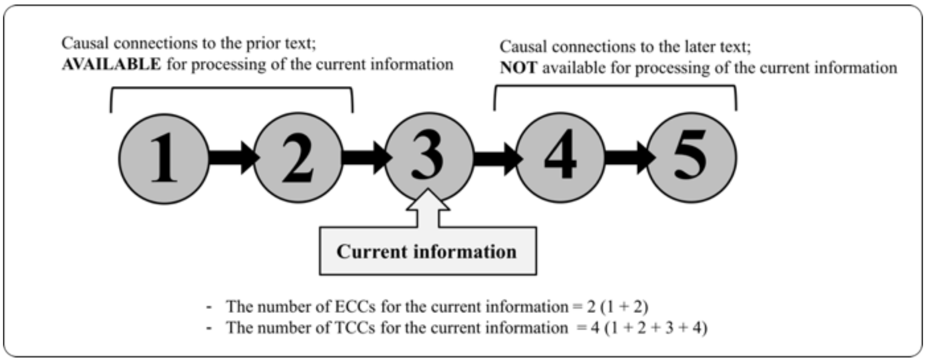

Figure 2.3. Two types of causal connections that are and are not available for processing of the current information.

Because past L2 studies exclusively employed TCCs, a theoretically sound account of how L2 readers process text's causal relations is not available.

Available information about how L2 readers process and encode scientific causal relations in expository text is even more limited. One of the few studies that applied the causal network model to L2 expository text comprehension was conducted by Ushiro et al. (2015). In their study, Japanese EFL readers' recall rates were found to increase linearly, with statements with the highest numbers of TCCs recalled best, followed by statements with moderate numbers of TCCs. Statements with the lowest numbers of TCCs were recalled worst. They interpreted this finding as showing that EFL readers' text memory reflected the causal structure of the expository text. However, they did not examine on-line processes. Whether EFL readers process causal relations in expository text as described by the causal network model remains unaddressed.

Although there has been no extension of the causal network model to L2 expository reading processes, causal bridging inferences were investigated by Hosoda (2014) with Japanese EFL readers of expository texts. He followed the methodology of past L1 expository studies (e.g., Noordman et al., 1992; Singer et al., 1997), using *because* target sentences (e.g., "Beginners are frequently advised to use compact skis because they usually have difficulty in changing directions") that required the generation of causal bridging inferences (e.g., "it is easier to turn on compact skis"). The results did not show clear evidence of EFL readers' inference generation from expository texts. However, only readers with high L2 reading proficiency seemed to detect causal gaps in target sentences that needed to be filled by inferences; their reading times for target sentences took significantly longer in the condition where inference ideas were not stated than in the condition where inference ideas were explicitly stated. On the other hand, low-proficiency readers did not detect causal gaps in target sentences—their target reading times did not differ, regardless of the presence of inference ideas in the text. Hence, it is likely that causal bridging inference generation in EFL reading may be affected by L2 reading proficiency. However, Hosoda's study was limited in that he did not control the familiarity of text content, though content familiarity has been shown to be influential in inference generation from expository text (Singer et al., 1997). The conditions under which EFL readers generate causal bridging inferences while reading expository text have not been fully clarified.

2.3 Measures to Assess On-Line Processes, Off-Line Memory, and Situation Models

So far, I have reviewed theoretical views and empirical findings regarding L1 and L2 reading. In this section, I provide a review of the methodology employed in this study to assess EFL readers' on-line processes, off-line memory, and situation models.

2.3.1 On-line measures

2.3.1.1 Three-pronged approach

This study used a *three-pronged approach* to examine on-line expository reading processes in EFL readers (Magliano & Graesser, 1991; Magliano et al., 1999; Suh & Trabasso, 1993; Trabasso & Suh, 1993). In this approach, a pair of complementary on-line measures are employed to investigate potential reading processes pre-identified by a discourse model.

Specifically, the first prong was a theoretical model of discourse analysis. The discourse model was used to identify potential a priori reading processes. For this study, I employed the causal network model to specify a set of potential causal bridging processes implemented during reading (e.g., Millis et al., 2006; Trabasso et al., 1989). The second prong was an unobtrusive behavioral on-line measure to examine whether readers engage in the processes predicted by the discourse model. I employed the *self-paced reading* method to this end. The third prong comprised verbal protocols that give an account of contents of processes, as informed by the second prong. I employed the think-aloud method. In sum, the three-pronged approach collected convergent evidence from two mutually complementary on-line measures (i.e., the self-paced reading and think-aloud methods) to elucidate whether one implements reading processes that are predicted by discourse analysis (i.e., the causal network model). Because the causal network model has already been described in Section 2.1.3, the next parts focus on the self-paced reading and think-aloud methods.

Self-paced reading method. The self-paced reading method is one of the simplest and most widely used measures to psychologically investigate on-line processes. The self-paced reading method usually involves recording participants' reading times for a pre-determined segment (e.g., a word, a phrase, a clause, or a whole sentence) of text. The theoretical rationale for the self-paced reading method is the *eye-mind assumption* (Just & Carpenter, 1992). According to the eye-mind assumption, the amount of time taken to read any given information indicates the amount of attention needed to understand that information. Therefore, reading times are assumed to become longer for text where participants make inferences to derive implicit information (e.g., Singer et al., 1997) or where they experience comprehension difficulty associated with conceptual gaps or textual inconsistencies (e.g., O'Brien et al., 1998). On the other hand, shorter reading times are interpreted as indicating that participants' processing is facilitated (e.g., Radvansky et al., 2014).

An advantage of the self-paced reading method is that it is unobtrusive. Researchers do not have to interrupt participants' reading behavior to collect data. Thus, the method can offer a "pure" account of the time course of on-line processes.

At the same time, the self-paced reading method is limited, in that only temporal data (i.e., time taken to understand any given text segment) are available. In other words, this measure does not provide specific information about what processes readers engage in at a specific time point of reading. To complement this shortcoming, the present study's three-pronged approach employed the think-aloud method.

Think-aloud method. The think-aloud method is a widely used on-line measure to assess cognitive processes or strategies available to readers during comprehension (see Ericsson & Simon, 1993, for a review). In the think-aloud method, readers are asked to verbalize whatever thoughts come to mind after comprehending a designated segment of text. Verbal protocols obtained from readers thinking aloud can offer information about what is consciously available in working memory and codable in language when understanding target text segments.

One of the major benefits of the think-aloud method is that the data are rich in content. Verbal protocols can provide direct insights into the actual processes occurring at any time point of reading, which can overcome the self-paced reading method's exclusive reliance on temporal data. The richness of verbal protocol data then allows researchers to compare different types of processes in terms of (a) the extent to which they are implemented while reading (Trabasso & Magliano, 1996), (b) the amount of attention allocated to them by readers (e.g., Horiba, 2000; McNamara, 2004), (c) the relative difficulty in their implementation (Magliano & Millis, 2003; Shimizu, 2015), and (d) their relations to off-line memory (Magliano et al., 1999; Trabasso & Suh, 1993). Despite these advantages, the think-aloud method has several problems. First, what can be assessed is limited to what readers can consciously access and is codable in language. In other words, think-aloud protocols may not be informative about processes that occur passively or automatically during comprehension (e.g., word recognition in skilled readers). Second, the very instruction to verbalize thoughts might encourage readers to take a more active reading strategy than usual, which alters the quality of processes (i.e., *reactivity effects*). The critical point is that both of these problems are attributable to an aspect of the think-aloud method that are not involved in normal reading; readers are required to translate their thoughts into language. To overcome this problem, I also used the self-paced reading method, which is assumed to be unintrusive. Therefore, employing the three-pronged approach in this study allowed for the collection of multifaceted complementary data on EFL readers' on-line processes.

2.3.1.2 Joint measures of target reading times and inference response times

To specifically assess the generation of causal bridging inferences during EFL expository reading, this study employed joint measures of inference processing developed by Noordman et al. (1992) and Singer et al. (1997). Table 2.4 shows an example passage used in this paradigm. This paradigm employed two measures to index readers' inference generation: reading times for target sentences and response times for inference questions.

Table 2.4

Two Measures of On-Line Causal Bridging Inference Generation

- The presentation order in the explicit condition：(a)→(b)→(c)→(d)→(e)→(f)
- The presentation order in the implicit condition：(a) → (c)→(d)→(e)→(f)

(a) Skiers of different abilities need different equipment from compact to large.

(b) It is easier to turn on the shorter compact skis.

(c) Beginners are frequently advised to use compact skis because they usually have difficulty in changing directions. (target)

(d) Once they can control their movement, they can quickly advance in skill.

(e) Is it easier to turn on compact skis? (inference question)

(f) Is it hard for novices to quickly advance in skill? (detail question)

Note. Reading and response times for the underlined parts are recorded.

The first index of inference generation comprised reading times for target sentences that require causal bridging inference generation for understanding (e.g., "Beginners are frequently advised to use compact skis because they usually have difficulty in changing directions"). If readers make causal bridging inferences during reading, reading times for target sentences should be significantly longer when inferences are implicit (implicit condition) than when they are explicitly stated (explicit condition) before target sentences. This prediction follows the assumption that inference generation takes additional time (Just & Carpenter, 1992).

The second index comprised response times for inference questions probing inference ideas in focus. With causal bridging inferences generated, readers should take comparable time to answer inference questions in the implicit and explicit conditions—if readers make inferences during reading, the corresponding ideas should be active in their memory, and inference questions in the implicit condition could be verified at a similar speed as in the explicit condition.

Note that neither of these two indices (target reading times, inference response times) are sufficiently sensitive to detect causal bridging inferences. Reading times do not provide direct information about actual processes, and response times cannot assess the time course of inference generation. However, combining these two measures allows for the simultaneous examination of the time course (as reflected by reading times) and contents (as reflected by response times) of inferences. Thus, like the three-pronged approach, these joint measures provide a complementary way of examining the moment-by-moment status of inference generation (Singer et al., 1997).

2.3.2 Off-line measures

I used a recall task and a causal question to assess participants' off-line memory and causal understanding of text, respectively.

2.3.2.1 Recall

The recall task (hereafter "recall" for simplicity) is probably the most widely used off-line measure to assess participants' off-line text memory. The assumption underlying recall is that information about which readers build mental representations is stored in memory and recallable after reading. Following this basic assumption, researchers interpret recall rates (the amount of information recalled / the total amount of information in text) as an index of how much information participants memorized from a text. Within the framework of the causal network model, recall rates of statements with different numbers of TCCs are compared to identify whether participants' off-line memory reflects the causal structure of the original text (i.e., higher recall rates of statements with many TCCs than statements with fewer TCCs are interpreted as evidence of memory reflecting causal structure; e.g., Varnhagen, 1991). Because instructions are simple and open-ended (i.e., participants are only asked to write down all information about a text), recall tasks impose few constraints or biases on participants' responses. Thus, recall provides a clearer window into off-line text memory than other measures (e.g., multiple-choice questions).

In addition, recall protocols can be analyzed qualitatively. The qualitative analysis of recall protocols provides information about how participants organize bits or pieces of text in memory (Coté et al., 1998). For example, Ushiro et al. (2015) examined how causally coherent text ideas were sequenced in recall protocols by Japanese EFL readers, aiming to qualitatively assess the internal coherence of memory representations. In the present study, I jointly employed quantitative and qualitative analyses of recall protocols to obtain a broader picture of EFL readers' memory for expository text.

Although recall provides useful quantitative and qualitative data about participants' text memory, several cautions must be noted. First, recall cannot directly assess situation models (i.e., learning from text), because it almost exclusively requires the reproduction of explicit text information. Hence, recall performance should be interpreted as an indication of textbase (i.e., memory for the explicit text), not situation models that go beyond the explicit text (e.g., McNamara et al., 1996). Second, what readers simply remember after reading does not necessarily equal what they understand from text (e.g., Kintsch, 1994). For example, readers can retrieve and recall individual text events from their working memory, even when they do not have a causal understanding of how those events are mutually interrelated in the described situation. Thus, to assess participants' causal understanding of expository text, this study employed *causal questions*.

2.3.2.2 Causal question

A causal question is a why-type question probing a series of causal relations in text (e.g., "Why does 'staying in zero gravity' lead to 'the heart becoming smaller'?" for an expository text discussing how zero-gravity environments affect the human body). Causal questions require readers to causally explain the queried text events (e.g., Carlson, van den Broek, McMaster, Rapp, Bohn-Gettler, Kendeou, & White, 2014). As mentioned in Section 2.1.3, causal explanation can be an indicator of readers' understanding of causal relations (Graesser & Clark, 1985; León & Peñalba, 2002). To causally explain a text, readers must coherently understand how and why a series of given text events lead to and result from other events (Chi, 2000; Coté et al., 1998). Therefore, by eliciting causal explanation, causal questions tap into readers' causal understanding more directly than the recall test, whose performance largely depends on memory for explicit text information.

At this point, it must be noted that previous L2 reading studies on causal relations exclusively used the recall test (e.g., Horiba, 1996; Ushiro et al., 2010: Ushiro et al., 2015) and did not use causal explanation to operationalize participants' causal understanding. To avoid any confusion, it must be re-emphasized that recall can be used to assess the causal structure in memory (i.e., whether statements with many TCCs are recalled better than statements with fewer TCCs), but can only partially assess understanding of relations between pieces of information. Therefore, how well L2 readers causally explain a text when they are given a causal question remains unaddressed.

In psychological terms, integrating pieces of relevant information into causal explanation requires not only memorizing each piece of relevant information but also interconnecting a series of information units in memory (McNamara & Kintsch, 1996).

This process can be explained in terms of the *fan effect* (Anderson & Reder, 1999). The fan effect is a phenomenon where activation spreads across all information units, reducing the activation for each one and decreasing its availability. When pieces of information are not interconnected, the fan effect may occur because there is no principled guide for activation (Reder & Anderson, 1980; Reder & Ross, 1983).

To avoid the fan effect, one must have relevant information connected in a way so that activation can converge on the information that is most likely to be the causal explanation (Millis & Barker, 1996). Thus, to understand the nature of EFL readers' causal explanation, it is necessary to examine how text memory's qualitative (how interrelated text information is memorized) and quantitative status (how much information is recalled) are related to performance on responding to the causal question.

2.3.3 Situation model measures

The construction of situation models involves readers going beyond explicit text (e.g., Kintsch, 1998). However, past L2 reading studies have exclusively used reproductive measures (recall, summary) focusing on explicit text (e.g., Horiba, 2000; Ushiro et al., 2015). Few L2 studies have used situation-model measures that directly assess L2 readers' comprehension beyond explicit text. To overcome this methodological limitation of past L2 studies, this study employed two situation-model measures. The first examined inferences in responses to off-line measures. The second was a *problem-solving test*, where participants were asked to apply learned information to a new situation.

2.3.3.1 Inferences in responses

This measure basically follows the notion that the extent to which off-line measures (e.g., comprehension questions, recall) indicate textbase or situation models is on a continuum, varying as a function of the amount of inferences included in responses; higher numbers of inferences are indicative of situation models (McNamara & Kintsch, 1996). In order to assess how well participants build situation models of causal relations, this study examined the amount of three types of inferences in answers to the causal question (Barry & Lazarte, 1998): (a) within-text inferences (logical interpretations from text), (b) elaborative inferences (combined propositions of text ideas and readers' knowledge relevant to the text topic), and (c) incorrect inferences (propositions that are contradictory or irrelevant to what is conveyed by text). Barry and Lazarte interpreted the total amount of these inferences as indicative of the richness of situation models, and proportions of incorrect inferences (incorrect inferences / [within-text + elaborative + incorrect inferences]) as indicative of the accuracy of situation models.

A main benefit of this measure is that it provides a qualitative account of situation models. For example, when one participant shows more incorrect inferences than another, researchers can assume that the former built situation models of lower accuracy than the latter. Still, it must be noted that this measure does not directly elicit participants' application of knowledge learned from text. In other words, how well a reader can use learned knowledge in a novel situation cannot be directly assessed. To overcome this limitation, I employed a problem-solving test.

2.3.3.2 Problem-solving test

The problem-solving test asks readers to apply learned information to a new situation to solve problems or explain phenomena (e.g., Mautone & Mayer, 2001). This study employed the problem-solving test used by McCrudden et al. (2009) and Mautone and Myer (2001). In these studies, L1 college students answered problem-solving questions such as "How could a space station be designed so that astronauts would be less likely to develop kidney stones?" after reading an expository text on how kidney stones develop in space. This process requires that both knowledge of the mechanism underlying the development of kidney stones (e.g., astronauts' bones become weaker, calcium levels in blood get higher, and the kidney needs to filter greater amounts of calcium from the blood) and knowledge about how to improve bone strength (e.g., doing physical exercises) are available in long-term memory. In other words, successful problem solving cannot be achieved unless readers make inferences from prior knowledge, as well as from the text in question.

It is important to note that the problem-solving test has been used in previous research on second-language acquisition and applied linguistics (Centeno‐Cortés & Jiménez, 2004; Robinson & Ross, 1996). The point is, however, that almost no L2 reading studies have used such applied measures, including the problem-solving test, to assess knowledge (i.e., information represented in situation models) that has been gained by L2 readers from text. To overcome this limitation, this study used the problem-solving test as a situation-model measure. The problem-solving test explicitly requires readers to go beyond the text. By using this test, I aimed to more directly assess EFL readers' ability to apply what has been learned from text than can be assessed with a reproductive measure.

2.4 Summary of the Previous Findings and Overview of the Present Research

I provided a theoretical, empirical, and methodological review of relevant literature in the previous sections. This final section of the literature review summarizes what remains unknown about causal understanding and learning from text in L2 reading. How the present study addressed those unresolved points is also discussed.

2.4.1 What remains unknown about causal understanding and text learning in EFL reading

(1) What on-line processes and off-line memory are involved in understanding of causal relations

First, there is no grounded account of how EFL readers understand causal relations in expository text. To be more specific, it is unclear what on-line reading processes and off-line text memory are involved in EFL readers' causal understanding of expository text.

The limited understanding of EFL readers' causal understanding is at least partly attributable to the absence of research exploring their causal explanation. In cognitive terms, causal understanding of text can be defined by the ability to causally explain the text (e.g., León & Peñalba, 2002; Millis et al., 2006). However, most previous L2 studies on causal relations have used the recall test, for which performance (recall rates) does not necessarily indicate relational understanding (Kintsch, 1994). Thus, there is no grounded empirical work on how EFL readers achieve causal understanding of text.

(2) How causal understanding and learning from text are related

Learning from text is construed as the construction of situation models, by which text information is integrated with readers' knowledge (e.g., Kintsch, 1994). Several L1 discourse models and empirical studies have reported that understanding causal relations in text is critical for situation-model construction (e.g., Linderholm et al., 2000; Millis et al., 2006).

Unlike for L1 readers, it is unclear how causal understanding in EFL readers relates to their final learning outcomes from text, partly because little L2 research has assessed situation models using appropriate methodology. Previous L2 reading studies have used reproductive measures (e.g., recall, summary) that can be completed with the original explicit text to assess situation models (e.g., Horiba, 2000; Ushiro et al., 2015). It remains unknown whether EFL readers productively apply learned knowledge from text to a new situation, and to what extent they do so.

(3) Under what condition EFL readers make causal bridging inferences during expository reading

Building situation models of text's causal relations involves the generation of causal bridging inferences (e.g., Noordman et al., 1992; Singer et al., 1997). However, most previous L2 inference studies were conducted using narrative texts (e.g., Muramoto, 2000; Yoshida, 2003). There is little direct evidence of the conditions under which EFL readers make causal bridging inferences during reading expository text.

Past L1 studies have reported that inference generation from expository text depends on the familiarity of text content (Noordman et al., 1997; Singer et al., 1997). When text content familiarity is high, reading processes are facilitated by the passive activation of prior knowledge, thus promoting inference generation (e.g., Gerrig, 2005; O'Brien et al., 1998). At the same time, the adequate use of content familiarity usually builds on the efficient operation of lower-level reading processes (Kintsch, 1998), which may become difficult when L2 readers' proficiency is low (Horiba, 1996). From this perspective, an interactive effect of the familiarity of text content and L2 reading proficiency may affect EFL readers' causal bridging inference generation.

(4) Whether, or to what extent, L2 readers' on-line processes reflect the causal structure of expository text

It has been shown that L2 readers generally build text memory reflecting text's causal structure, as described by the causal network model (Langston et al., 1999; Millis et al., 2006). Specifically, L2 readers tend to recall information with many TCCs better than information with fewer TCCs after reading expository (Ushiro et al., 2015) and narrative texts (Horiba, 1996; Ushiro et al., 2010).

However, there is no appropriate extension of the causal network model to L2 readers' on-line processes. Although several L2 narrative studies have used the causal network model to explore on-line processes (e.g., Horiba, 1996; Ushiro et al., 2010), they analyzed the total number of causal connections that any unit of information has to other information in the text (i.e., TCCs). This approach is problematic because, as pointed out by Radvansky et al. (2014), on-line processes should be affected by causal connections current information has to prior text, but not to later text. Therefore, we have limited understanding of whether, or to what extent, EFL readers' on-line expository reading captures causal structure as predicted by the causal network model.

2.4.2 How the present research addressed the limitations and the problems of past research

To address the unresolved points described above, I conducted two studies (Study 1, Study 2), each consisting of three experiments.

Study 1. Study 1, which consisted of Experiments 1, 2, and 3, examined EFL readers' understanding of causal relations and learning from text. Experiment 1 explored how EFL readers' quantitative (i.e., recall rates of text) and qualitative (i.e., internal coherence of recall protocols) text memory were related to their causal explanation elicited by a causal question. By doing so, I sought to reveal the nature of text memory involved in EFL readers' causal understanding of expository text. Experiment 2 explored the relation between EFL readers' causal understanding and their learning outcomes from expository text. I focused on L2 reading proficiency, which has been shown to affect L2 readers' situation-model construction (e.g., Muramoto, 2000; Shimizu, 2015). Based on the results from Experiment 2, Experiment 3 was conducted to gain a better understanding of text-learning difficulties experienced by low-proficiency EFL readers.

Study 2. Study 2, which consisted of Experiments 4, 5, and 6, examined on-line processes involved in EFL readers' expository text comprehension. Experiment 4 explored the conditions under which EFL readers generate causal bridging inferences during expository reading. I placed a special focus on interactive effects of L2 reading proficiency and familiarity of text content on inference generation. Experiments 5 and 6 employed the three-pronged approach, together with off-line measures, to examine how EFL readers' on-line processes and off-line memory reflect the causal structure of expository text, as identified by the causal network model. Experiment 5 used the self-paced reading method and the recall test, aiming to reveal whether, and to what extent, EFL readers' reading times for text statements and recall probability reflect a text's causal structure. Experiment 6 used the think-aloud method to examine contents of processes involved in EFL readers' expository comprehension. Experiment 6 also used the causal question and explored how EFL readers' off-line causal understanding and on-line causal bridging are related.

Finally, with the completion of Studies 1 and 2, I aimed to advance an understanding of EFL readers' expository text comprehension in terms of the three General RQs.

General RQ1 What on-line processes and off-line memory are involved in EFL readers' causal understanding of expository text? (Experiments 1, 6)

General RQ2 How does EFL readers' causal understanding contribute to their learning outcomes from expository text? (Experiments 2, 3, 4)

General RQ3 How do EFL readers' on-line processes and off-line memory reflect the causal structure of expository text? (Experiments 5, 6)

Chapter 3

Study 1: Causal Understanding and Learning From Expository Text in EFL Reading

3.1 Experiment 1

3.1.1 Purpose and research question

The purpose of Experiment 1 was to explore the relation between EFL readers' causal understanding of expository text and their off-line text memory. The overall goal of Study 1 was to explore the relation between EFL readers' understanding of causal relations and final learning outcomes from expository text. This experiment was designed as a starting point for this goal, seeking to explore EFL readers' causal understanding in terms of their text memory. Following previous L1 reading research (e.g., McCrudden et al., 2009; Millis et al., 2006), I operationalized understanding of causal relations as performance of causal explanation elicited by a causal question. Participants' off-line memory was assessed by a recall test.

Specifically, I was interested in whether and how the quantity and quality of text memory differ between readers who can and cannot causally explain expository text well. Researchers argue that integrating relevant text ideas into causal explanation requires that readers build efficiently organized text memory representations where bits of information are interconnected (Graesser & Hemphill, 1991; Millis & Barker, 1996). Otherwise, a fan effect would occur, in which activation haphazardly spreads across several concepts, making it difficult to identify relevant information for causal explanation. From this point of view, it is likely that readers have trouble using memorized information to form causal explanation without coherent connections between pieces of information.

Following this expectation, I qualitatively compared internal coherence of recall protocols by participants who answered the causal question well (termed *good explainers*) with that of participants who answered poorly (termed *poor explainers*), in addition to quantitatively analyzing intergroup differences in recall rates. The following research question (RQ) was addressed in this experiment:

RQ1: How do the quantity and the quality of text memory differ between EFL readers who can and cannot causally explain expository text well?

3.1.2 Method

3.1.2.1 Participants

Participants were 52 Japanese undergraduate and graduate students at University of Tsukuba and National Institute of Technology, Ibaraki College. Thirty were female, and 22 were male. Their ages ranged from 18 to 26 years old (M = 21.01, SD = 1.08). Their study majors were various, including biology, humanities, mathematics, psychology, education, and international studies. They all had studied Japanese for at least six years as part of the compulsory education in Japan. They all had lived in Japan for more than 18 years. None of them had studied abroad for more than two weeks.

According to their self-reports, participants' overall English proficiency was estimated to be beginner-intermediate to intermediate-advanced, or CEFR (Common European Framework of Reference for Languages) levels A2 to B2 (Council of Europe, 2001; Tannenbaum & Wylie, 2008, 2013); the TOIEC listening and reading test (M = 502.60, SD = 99.42, range = 330 to 900) and the EIKEN test (Grade 4 to Grade 1: Grade 4, n = 1; Grade 3, n = 7; Grade Pre-2, n = 7; Grade 2, n = 8; Grade Pre-1, n = 1; Grade 1, n = 0). These self-reports were collected by a paper questionnaire. Note that thirty-nine participants reported at least one of the scores, and the other 13 participants reported none of the scores.

3.1.2.2 Material

Text. An expository text discussing how staying in zero gravity causes astronauts' hearts to shrink was used (presented in Table 3.1). This passage was originally part of a longer passage used in McCrudden et al.'s (2009) experiment with L1 readers. Later, Ushiro et al. (2015) shortened the passage in their experiment with L2 readers. The present passage is an adapted version of Ushiro et al.'s.

Table 3.1

An Experimental Text Used in Experiments 2 and 3

When people first considered space travel, they did not know how the zero gravity of space would affect humans. In fact, the human body is a complex system that automatically responds to the lack of gravity.

While in space, the body is not affected by gravity. Therefore, blood and water do not travel to the lower parts of the body, especially the legs. Instead, the blood and water within the body move to the upper body. Because the blood and water travel to the upper parts of the body, the body feels like the chest and head are filled with blood and water. Because of this, the heart and lungs send messages that the amount of blood and water in the upper part of the body must be reduced. As a result, space travelers do not feel thirsty, and therefore, space travelers drink less water. As body water is eliminated, their body water levels become lower than normal. When the amount of blood and water decrease, it becomes more difficult for the human body to work normally. In addition, the decreased body water makes the heart pump less blood than normal. Therefore, the heart does not need to work as hard as it does on Earth. As a result, the heart becomes smaller.

Studying the effects of space travel on humans can help us better understand many illnesses, such as high blood pressure and other heart problems.

Specifically, I added some statements to Ushiro et al.'s (2015) passage to make clearer the discourse flow in the later part of the passage (why the weakening of the heart functioning leads to the heart shrinkage). I also deleted statements that were not relevant to the main theme of the passage to clarify the cause-effect sequences. Further, words assumed to be unfamiliar to EFL readers were replaced with more familiar words (e.g., fluids → water). This was done with reference to the *JACET (the Japanese Association of College English Teachers*) *list of 8000 basic words* (Ishikawa et al., 2003). This lists English words that Japanese students are supposed to learn from elementary school to university based on frequency from level 1 (most frequent) to 7 (least frequent). Words from level 5 or higher were rephrased with more frequent words from level 4 or below. After this, a native speaker of English checked the text for naturalness of expression and discourse.

Causal question. The causal question asked participants to causally explain why "staying in zero gravity" leads to "the heart shrinking," the series of causal relations described in the text. Appendix A shows the causal question and the instruction. Expected responses included a maximum of six causal relations (CRs; CR2–CR7), which are presented in Table 3.2. Note that the beginning (CR1 "Lack of Gravity") and the end (CR8 "Heart shrinks.") of the causal chain were provided as a cue on the answer sheet.

Table 3.2

List of Key Causal Relations (CRs) in the Experimental Text

CR1: Lack of Gravity
CR2: Body fluids shift headward.
CR3: Body senses flood of fluids.
CR4: Body tries to decrease fluids.
CR5: Body eliminates more fluids and consumes less fluid
CR6: Body fluid levels are decreased.
CR7: Heart does not work normally.
CR8: Heart shrinks.

The instruction for the causal question was as follows: “Why does ‘staying in zero gravity’ lead to ‘the heart becoming smaller’? Explain as much as possible to link these two events in a logically and causally correct order.” As in the instruction, participants were told that it was important to provide as many relevant causal relations as possible, which aimed to avoid confounding with the omission of understood relations from an answer. The instruction was given on the answer sheet, in Japanese (participants’ L1), and participants were also asked to answer in Japanese so that L2 writing skills would not affect the result. They were not allowed to refer to the text during the task.

3.1.2.3 Procedure

The experimenter explained the general purpose and procedures of the experiment to participants, and informed consent was obtained. Then, participants were given the experimental text on a sheet. They were asked to read the text for understanding. All participants completed reading within five minutes. The experimenter then took the text away from the participants so that they could not refer to it, and the recall test was administered. Participants were asked to write down all that they could remember from the text in Japanese. The recall test was finished within 15 minutes. The causal question was next administered. Participants were asked to explain why “staying in zero gravity” leads to “the heart becoming smaller” in as detailed a manner as possible, and in a causally correct order, in Japanese. The causal question was finished within 15 minutes. The total time to complete the experiment was approximately 60 minutes.

3.1.2.4 Scoring and data analysis

Recall. For quantitative analysis, two Japanese graduate students (including I) first divided the experimental passage into 29 statements, each corresponding to a subject-verb clause (Ushiro et al., 2010; Ushiro et al., 2015). Inter-rater agreement was 99%, with disagreements resolved through discussion. Five statements corresponding to the last paragraph were excluded from scoring to avoid any recency effect on recall. Two Japanese graduate students (including I) independently scored 30% of the data. A point was given if the gist of the statement (about two-thirds of the meaning) was recalled, but scientifically incorrect responses were not credited. Inter-rater agreement was 92%, and all disagreements were resolved through discussion. Using the refined criteria, I scored the remaining data.

In addition, I qualitatively examined the organization of recall protocols to assess the coherence of participants’ text memory. The following three criteria, constructed based on literature on memory networks, were used in scoring (Coté et al., 1998; Ushiro et al., 2015): (a) ideas leading to the direct cause of heart shrinkage were provided, (b) information was sequenced in a causally and logically correct order, and (c) statements were cohesively and meaningfully interrelated. If all criteria were satisfied, participants’ recall was classified as "+Coherence," meaning that they had built text memory with sufficient coherence. When participants failed to fulfill at least one criterion, their recall was classified as "−Coherence," meaning that their memory for the text lacked coherence. Two trained Japanese graduate students (including I) scored 30% of the data. First, we reached the inter-rater agreement of 80%. We had several discussions and repeated scoring until we reached the inter-rater agreement of 90%. With the refined criteria, I scored the remaining data.

Causal question. Understanding of causal relations was assessed in terms of the number of key causal relations (presented in Table 3.2) produced in the answers. Specifically, the answer scored 0 to 6 points depending on the number of causal relations produced in a causally correct order. As long as the order was correct, points were given even when intermediate relations between them were absent. However, answers that were scientifically incorrect, inconsistent with the text, or produced in an incorrect order were not credited.

For example, consider the answer, "When a space traveler stays in zero gravity, his body water goes up (CR2), and his body feels a lot of water (CR3). So, the heart does not work hard (CR7), and the water level then decreases (CR6)." This receives 3 points (CR2 + CR3 + CR7); CR6 does not earn a point because it comes after CR7, which should occur earlier. Two judges independently scored 30% of the data, resulting in inter-rater agreement of 90%. After disagreements were resolved through discussion, the author scored the remaining data.

3.1.3 Results

3.1.3.1 Descriptive statistics

Table 3.3 presents the means, standard deviations, and 95% confidence intervals (CIs) of recall rates and performance on the causal question. Based on causal-question performance, participants were grouped into poor and good explainers; those who scored 3 or above constituted good explainers (n = 28), and those who scored 2 or below constituted poor explainers (n = 24). An independent t test confirmed performance differences between the groups, $t(50) = 11.08$, $p < .001$, $d = 3.12$. In the following section, I quantitatively and qualitatively compare recall performance of these two groups to address the research question.

Table 3.3

Recall Rates and Causal-Question Performance in Poor and Good Explainers

	Poor explainers (n = 24)			Good explainers (n = 28)		
	M	*SD*	95% CI	*M*	*SD*	95% CI
Recall rates	.37	.17	[.31, .44]	.44	.14	[.39, .49]
Causal question	1.62	0.49	[1.43, 1.82]	4.07	0.98	[3.71, 4.43]

Note. CI = confidence interval.

3.1.3.2 Differences in text memory between good and poor explainers

First, to quantitatively compare text memory in good and poor explainers, recall rates were submitted to an independent t test. The results showed that good and poor explainers' recall rates were not significantly different, $t(50) = 1.63$, $p = .109$, $d = 0.45$, meaning that they did not vary significantly in terms of the amount of information they recalled from the text, as shown in Figure 3.1.

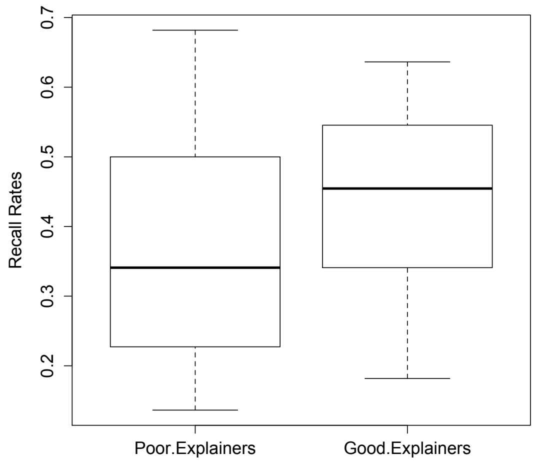

Figure 3.1. Recall rates in poor and good explainers.

Next, I compared good and poor explainers' internal coherence of recall protocols to examine whether text memory was qualitatively different between the two groups. The results are shown in Figure 3.2. It was found that the majority of good explainers produced recall protocols evaluated as +Coherence (58% [16 out of 28]). On the other hand, recall by most poor explainers was evaluated as −Coherence (75% [18 out of 24]). The intergroup difference was statistically significant, $\chi^2(1) = 5.47$, $p = .019$, $\varphi = 0.32$.

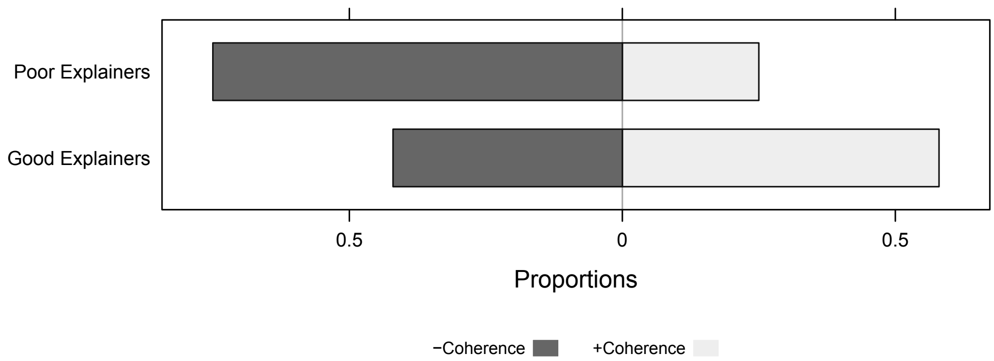

Figure 3.2. Proportions of participants who produced +Coherence and –Coherence recall in poor and good explainers.

Together, these results indicate that good explainers demonstrated qualitatively better recall than poor explainers.

3.1.4 Discussion

The results from the quantitative and qualitative analyses of recall indicate that text memory in participants who did (i.e., good explainers) and did not build causal understanding (i.e., poor explainers) differed qualitatively (i.e., how coherently information is interconnected in memory) rather than quantitatively (i.e., how much information is memorized from the text). Specifically, most participants who causally explained the text well were found to produce recall with coherent organization, indicating that they tended to build text memory where pieces of information were interconnected. On the other hand, recall by participants who did not causally explain the text well mostly lacked such internal coherence, meaning that they memorized text ideas in a relatively disconnected manner.

The fact that participants with and without causal understanding were discriminated by the quality of text memory is theoretically consistent with the fan effect (e.g., Anderson & Reder, 1999). This effect holds that converging the activation on relevant information becomes difficult when ideas are not stored in a mutually interconnected way. It can be assumed that the qualitative inefficiency in text memory posed difficulty in appropriately using memorized information to answer the causal question.

It was also revealed that the amount of information recalled from the text (i.e., recall rates) was not significantly different between participants with and without causal understanding. This implies that differences in causal-question performance cannot be sufficiently explained by the amount of information recalled from the text. The comparable recall rates between the groups may be attributable to the fact that the experimental passage was linguistically simplified. Because the linguistic difficulty and complexity of the passage was highly controlled, just encoding explicitly stated information into memory was generally easy for participants. Indeed, total recall rates were generally high (about 40%) for participants; in past L2 studies, total recall rates were generally 20% to 40% (e.g., Horiba, 2013; Ushiro et al., 2010; Ushiro et al., 2015).

Conclusively, combining qualitative and quantitative findings showed that text memory in participants without causal understanding was significantly less coherent than participants with causal understanding (as revealed by qualitative analysis), despite the fact that the former recalled as much information from the text as the latter (as revealed by quantitative analysis). These converging findings mean that the quantity of text memory does not necessarily guarantee the quality of it and the subsequent application of the memorized information to causal explanation. This claim is at least partly consistent with findings reported by Ushiro et al. (2015). In the study, Japanese EFL readers recalled statements with many causal connections after reading, regardless of L2 reading proficiency. However, participants in the lowest-proficiency group were found to produce recall that included incorrect cause-effect sequences or lacked interrelations among events. The present experimental findings provide evidence that such coherence in memory representations is critical for EFL readers' ability to use memorized text information to form causal explanation.

At the same time, it must be noted that it is unlikely that the amount of information memorized from text is always similar between those who can and cannot understand causal relations in text. Rather, I must admit that the observed comparable recall rates across participants with and without causal understanding were, to no small extent, due to the simplified nature of the experimental passage. From this perspective, it is safe to interpret the present results as showing that even when the amount of information memorized does not differ greatly, readers can have more difficulty with causal understanding when information is not interconnected in memory than when it is. This interpretation supports the notion that the understanding of causal relations cannot be operationalized by the amount of information recalled (Kintsch, 1994).

3.1.5 Conclusion of Experiment 1

The goal of Experiment 1 was to explore the relation between EFL readers' causal understanding and their memory for expository text. The results showed that text memory in good and poor explainers was qualitatively different. Good explainers produced coherent recall protocols significantly more often than poor explainers. On the other hand, the amount of information recalled was not significantly different between good and poor explainers.

These findings lead to the conclusion that just memorizing relevant information after reading is not sufficient for causal understanding. Instead, causal understanding involves readers building networked memory representations, where bits of relevant information are coherently interconnected.

I must note several unresolved points of Experiment 1. First, Experiment 1 did not consider participants' L2 reading proficiency, despite the fact that L2 reading proficiency has been reported to have an influence on L2 readers' comprehension (Muramoto, 2000; Ushiro et al.,2015). It is likely that EFL readers with different L2 reading proficiency build causal understanding to a different degree. Second, Experiment 1 did not explore participants' final learning outcomes from text; whether and how causal understanding of expository text allows EFL readers to learn text information as knowledge still remains unclear. To address these points, I designed Experiment 2..

3.2 Experiment 2

3.2.1 Purpose and research question

The purpose of Experiment 2 was to explore the relation between understanding of causal relations and learning outcomes from the expository text in EFL readers. Based on the notion that inferential processing in L2 readers is constrained by L2 reading proficiency (e.g., Hosoda, 2014; Yoshida, 2003; Li & Kirby, 2014), I considered readers' L2 reading proficiency as a variable. Specifically, past L2 reading studies have widely proposed that less proficient L2 readers have difficulty engaging in inference generation during reading (e.g., Horiba, 2000; Muramoto, 2000). Given the general agreement that inference generation is necessary for the construction of situation models (Kintsch, 1994, 1998; McNamara et al., 1996), it is likely that low-proficiency readers struggle to build situation models of text causal relations, and as a result, their causal understanding was expected not to contribute to text learning.

Participants' understanding of causal relations was assessed by the causal question ("Why does staying in zero gravity lead to the heart becoming smaller?"). This why-type question is widely used to elicit readers' causal explanation (e.g., Millis et al., 2006; Ushiro et al., 2015). To causally explain the text, readers must understand a series of causal relations between pieces of text information in an integrative manner (León & Peñalba, 2002). Thus, the causal question can be an indicator of causal relations understood and memorized by readers from the text. Furthermore, I inspected inferences in participants' causal question answers to examine how well participants constructed situation models of causal relations in the text (Barry & Lazarte, 1998).

For learning outcomes from the text, I used a problem-solving test to assess readers' application of knowledge gained from the text. Further, to obtain a detailed picture of the L2 reading proficiency effect on the problem solving, I qualitatively examined contents of problem-solving responses and identified patterns that were distinctive of low-proficiency readers. The following research question (RQ) were addressed in this experiment.

RQ2: Do contributions of EFL readers' understanding of causal relations to their learning outcomes from the text differ as a function of L2 reading proficiency?

3.2.2 Method

3.2.2.1 Participants

Participants were 70 Japanese undergraduate and graduate students at University of Tsukuba and National Institute of Technology, Ibaraki College. They were majoring in the humanities, education, medicine, physics, or psychology; 42 were female and 28 were male. None of them participated in Experiment 1. Their ages ranged from 18 to 24 years old (M = 20.01, SD = 3.16). They were all native speakers of Japanese and had lived in Japan for more than 18 years. They had studied English at least for six years as part of the compulsory education in Japan. None of them had experience of studying abroad for more than two weeks.

Based on their self-reports, participants' English proficiency was estimated to be from the CEFR A2 to B2 levels (Council of Europe, 2001; Tannenbaum & Wylie, 2008, 2013); the TOIEC listening and reading test (M = 560.94, SD = 100.03, range = 345 to 890) and the EIKEN test (Grade 4 to Grade 1: Grade 4, n = 2; Grade 3, n = 9; Grade Pre-2, n = 9; Grade 2, n = 13; Grade Pre-1, n = 2; Grade 1, n = 0). The self-reports were collected by a paper questionnaire. Fifty-five participants reported at least one of the scores, and the other 15 participants reported none of the scores.

Participants were grouped by means of a median split (*Mdn* = 9) into two proficiency groups based on their performance on an L2 reading proficiency test (see Table 3.4; *M* = 10.43, *SD* = 5.85, *Min*/*Max* = 3/26). Participants who scored lower than the median formed a *low-proficiency group* (*n* = 34, *M* = 5.94, *SD* = 1.77, *Min*/*Max* = 3/8). The low-proficiency group was estimated to be placed in the CEFR A2 level, according to their self-reports (the TOIEC listening and reading test: *M* = 400.87, *SD* = 72.65, range = 345 to 600; Tannenbaum & Wylie, 2008, 2013). On the other hand, participants who scored higher than the median formed an *intermediate-high-proficiency group* (*n* = 36, *M* = 14.67, *SD* = 5.15, *Min*/*Max* = 9/26). The intermediate-high-proficiency group was estimated to be placed in the upper end of the CEFR B1 level (the TOIEC listening and reading test: *M* = 680.12, *SD* = 50.90, range = 500 to 890; Tannenbaum & Wylie, 2008, 2013), and was named accordingly.

There was a significant difference in proficiency-test scores between the two groups, $t(68) = 9.6$, $p < .01$, $d = 2.26$. Table 3.5 the estimated CEFR levels and proficiency test scores of participants in Experiments 2–6.

Table 3.4

*L2 Reading Proficiency Scores of Participants in Experiment 2 (*N *= 70)*

Proficiency	*M*	95% CI	*SD*	Min	Max
LP (*n* = 34)	5.94	[5.34, 6.54]	1.77	3	8
IHP (*n* = 36)	14.67	[12.98, 16.35]	5.15	9	26

Note. LP = low proficiency. IHP = intermediate-high proficiency. CI = confidence interval. Scores on the L2 reading proficiency test range from 0 to 26.

Table 3.5

Experiments 2–5's Proficiency Groups With L2 Reading Proficiency Test Scores and the Estimated CEFR Levels

	Proficiency group							
	Low (the CEFR A2)		Intermediate-low (the lower end of the CEFR B1)		Intermediate-high (the upper end of the CEFR B1)		High (the CEFR B2)	
	L2 reading proficiency test scores							
Experiment	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Experiment 2	5.94	1.77			14.67	5.15		
Experiment 3	5.24	1.48			12.80	3.70		
Experiment 4			9.48	2.86			17.52	3.38
Experiment 5			8.65	2.10			16.28	2.79
Experiment 6			7.97	2.47			18.79	3.27

3.2.2.2 Materials

L2 reading proficiency test. The L2 reading proficiency test consisted of 26 items. These were derived from the reading section of the second (k = 20) and pre-first grades (k = 6) of the EIKEN Test in Practical English Proficiency (Obunsha, 2005a, Obunsha 2005b). The EIKEN test is one of the most widely used standardized English proficiency tests among Japanese educators and researchers (e.g., Hosoda, 2014; Ushiro et al., 2015). The 26 items of the present proficiency test tap into understanding of the content of passages rather than specific lexical or grammatical knowledge. Therefore, this proficiency test is assumed to target ability to comprehend English text appropriately and efficiently. The reliability of the test was acceptable, with Cronbach's α = .81.

Text. The same expository text was used as in Experiment 1.

Causal question. The same causal question was used as in Experiment 1.

Problem-solving test. The problem-solving test consisted of four questions requiring participants to use the scientific principles or mechanism learned from the text to solve problems or explain situations outside the original text. An example question and its expected response are presented in Table 3.6. As seen there, correct responses should not only be consistent with the text but also include inferences from the knowledge gained from the text.

The four problem-solving questions were prepared through a pilot study. Six candidate questions were first constructed through discussion between the author and a graduate student majoring in English education based on three criteria: (a) correct responses require readers to make inferences from the text, but (b) do not demand technical knowledge, and (c) queried situations are not restricted to the context of the text (McNamara & Kintsch, 1996). Forty-three Japanese EFL students read the experimental text and answered the created six candidate questions. The results indicated that participants' responses to two of the six questions varied excessively. Thus, they were excluded, and the remaining four questions were retained.

Table 3.6

An Example Problem-Solving Question and Expected Response (Translated From Japanese)

Question:

Imagine that you got a disease that makes you insensitive to changes in body fluids. Explain why, in this condition, your heart size would not change even if you stayed in space.

Expected Response:

(If the body were insensitive to changes in fluids), the body cannot eliminate fluids even though the blood and water collect in the upper body in space. So, the amount of body fluids would not be reduced, which causes the heart to pump as strongly as on the earth. Because the heart works as usual, the heart size would not change.

Appendix B presents all the four problem-solving questions and the instruction. These questions were presented to participants in Japanese in a fixed order. Participants were asked to write down their answers to the questions in as much detail as possible in Japanese. Participants were also asked to provide reasons for each answer based on what they had learned from the text.

3.2.2.3 Procedure

The experiment consisted of the first and the second (one week after text reading) sessions. It is important to note that the problem-solving test needed to be done in the second session because the intention was to assess learning from the text, not immediate memory for it.

First session. The experimenter explained the general purpose and procedures of the experiment to participants, and informed consent was obtained. Then, participants were given the experimental text on a sheet. They were asked to read the text for understanding at their own pace and turn the sheet back after reading the text once. There was no time limit for the reading. All the participants finished within five minutes. The experimenter then took the text away from the participants so that they could not refer to it any more. The causal question was then administered. Participants were asked to explain why ‘staying in zero gravity’ leads to ‘the heart becoming smaller’ in as detailed a manner as possible in a causally correct order in Japanese. There was no time limit for the causal question either. All the participants completed the causal question within 15 minutes.

Second session. One week after the first session, participants were gathered, and the problem-solving test was administered. There was no time limit on the problem-solving test. They were asked to write down their answers as detailed as possible in Japanese based on what they had learned from the text (McCrudden et al., 2009). They were not allowed to refer to the text during the task. The problem-solving test was finished within 20 minutes. Finally, an L2 reading proficiency test was administered for 30 minutes. The total time to complete the experiment was approximately 90 minutes.

3.2.2.4 Scoring and data analysis

Causal question. Answers to the causal question was scored in the same way as in Experiment 1. Two trained raters scored 30% of the data. The inter-rater agreement was 95%. All the disagreements were resolved through discussion. Using the refined criteria, the author scored the remaining data.

As for inferences in responses, two judges first counted inferential ideas found in causal question answers through discussion. The unit of analysis was a subject-verb clause in Japanese. The found inferences were subsequently classified into the following three categories (Barry & Lazarte, 1998): (a) within-text inferences, (b) elaborative inferences, and (c) incorrect inferences. Within-text inferences indicated readers' building coherence across pieces of the text. This was often associated with adding of explanatory information to the explicit text (e.g., the underlined part of "The muscle in the heart is reduced because it does not need to work hard"). Elaborative inferences indicated readers' embellishing mental representations with prior knowledge. This was also associated with adding relevant information to text ideas, but it did not contribute to the coherence (e.g., "While in space, an astronaut's face will swell because fluids shift to the upper body"). Incorrect inferences indicated readers' misunderstanding or use of irrelevant information to the text. This type of inference included off-topic information and ideas contradictory to what was mentioned in the text (e.g., "The heart sends greater amounts of blood than usual as the body fluids decrease in space"). The two judges independently categorized 30% of the inferences, resulting in inter-rater agreement of 90%. After disagreements were resolved through discussion, the author categorized the remaining 70% of data.

Problem-solving test. Participants' responses were assessed by a scoring system developed in past expository comprehension research (Gilliam et al., 2007; Magliano et al., 2011; Millis et al., 2006). This identified necessary information for the pre-created ideal answers for each question. The ideal answers in this study were constructed in a four-step procedure. First, four judges (all familiar with the experimental text) independently created candidate answers to the four problem-solving questions. Second, these candidate answers were parsed into individual ideas (Japanese clauses). Third, the four judges held a discussion to identify ideas that were necessary for each answer to the problem-solving questions; ideas were deemed as necessary when three or more of the judges included in their candidate answers. Fourth and finally, the ideal answers were determined by assembling the identified necessary ideas.

Using the ideal answers, responses were scored using a 4-point scale (0–3): 0 meant that the answer was incorrect; 1 meant that the answer was vague but correct on the whole; 2 meant that the answer was partially complete; 3 meant that the answer was complete. The total scores ranged 0 to 12. Two judges scored 30% of the data. The inter-rater agreement was 93% with disagreements resolved through discussion. The author scored the remaining data.

For the qualitative analysis on the problem-solving responses, two judges first separately examined the responses, and a list of response patterns that were distinctive of the low-proficiency group was constructed. Next, the two judges had a discussion and eliminated those patterns from the list that were not agreed on by the two raters and that overlapped with other patterns. Then, proportions of participants showing the selected patterns were compared between the proficiency groups, using chi-squared tests. Finally, patterns whose p values were below .05 were deemed as distinctive of the low-proficiency group.

3.2.3 Results

The results section consists of the following four parts. First, I report descriptive statistics of the causal question and the problem-solving performance. Second, contributions of causal question performance to the problem solving are reported. Third, results regarding inferences in the causal question are reported. Finally, I report distinctive patterns of problem solving in the low-proficiency group.

3.2.3.1 Descriptive statistics and intergroup differences

Table 3.7 presents means and standard deviations of the two proficiency groups' performance on the causal question and the problem-solving test. For the causal question, the intermediate-high-proficiency group performed significantly better than the low-proficiency group, $t(68) = 2.94$, $p = .004$, $d = 0.70$, as shown in Figure 3.3.

Table 3.7

Performance on the Causal Question and the Problem-Solving Test in Experiment 2

	Causal question			Problem solving		
Proficiency	*M*	95%CI	*SD*	*M*	95%CI	*SD*
LP (*n* = 34)	1.85	[1.46, 2.24]	1.16	2.88	[2.32, 3.45]	1.68
IHP (*n* = 36)	2.75	[2.30, 3.20]	1.38	5.47	[4.41, 6.54]	3.26

Note. LP = low proficiency; IHP = intermediate-high proficiency; Scores on the causal question and the problem-solving test range 0 to 6 and 0 to 12, respectively.

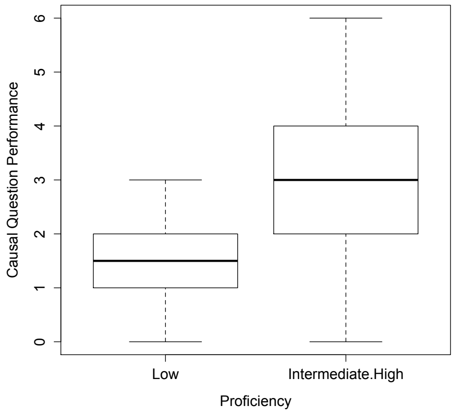

Figure 3.3. Causal-question performance in Experiment 2.

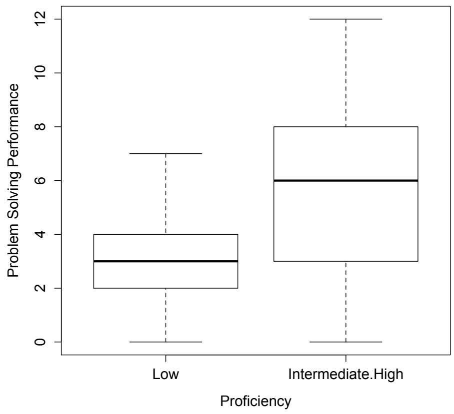

Figure 3.4. Problem-solving performance in Experiment 2.

Specifically, the majority of the low-proficiency group (79% [27 out of 34]) recalled only two or less of the six target causal relations. Similarly, the problem-solving performance was significantly better in the intermediate-high- than the low-proficiency group, $t(68) = 4.14$, $p < .001$, $d = 0.99$ (see Figure 3.4). Hence, it was confirmed that the intermediate-high-proficiency group outperformed the low-proficiency one in both understanding of causal relations and learning outcomes from the text.

3.2.3.2 Contributions of causal-question performance to the problem solving

To examine the relation between understanding of causal relations and text learning, I first computed correlations between causal-question and problem-solving performance separately for the two proficiency groups. The results revealed quite different pictures between the groups (see Figures 3.5, 3.6).

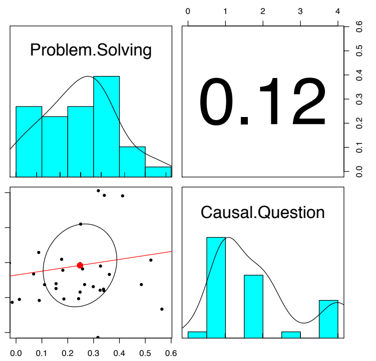

Figure 3.5. The scatter plot of the correlation between performance on the causal question and the problem-solving test in the low-proficiency group in Experiment 2.

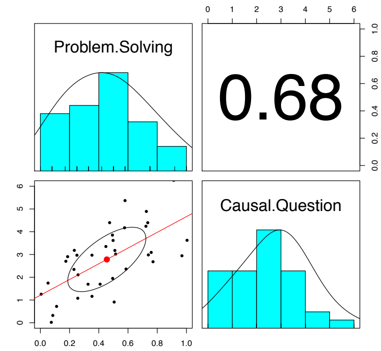

Figure 3.6. The scatter plot of the correlation between performance on the causal question and the problem-solving test in the intermediate-high-proficiency group in Experiment 2.

In the intermediate-high-proficiency group, performance on the causal question and the problem-solving test were significantly correlated, $r = 67$, $p < .001$. By contrast, this correlation in the low-proficiency group was much lower and failed to reach significance, $r = 12$, $p = .404$. This intergroup difference was statistically significant, $z = 2.68$, $p = .007$.

Based on the correlation results, I next ran a hierarchical regression analysis using the problem-solving performance as a dependent variable to clarify whether L2 reading proficiency modified contributions of causal question performance to the problem solving. In Steps 1 and 2, L2 reading proficiency test scores (termed Proficiency) and causal question performance (Causal Understanding) were entered as predictor variables, respectively. In Step 3, I entered the interaction term of proficiency and causal understanding (the Proficiency × Causal Understanding interaction), which was created by multiplying L2 reading proficiency test scores by causal-question performance. The focus was on the significance of the R^2 change associated with the entry of the Proficiency × Causal Understanding interaction; this value represented the extent to which the interaction explained the variance of the problem-solving performance above and beyond the main effects of proficiency and causal understanding. The results did confirm the Proficiency × Causal Understanding interaction, β = .26, p = .015, with its entry increasing the model's predictive power by 5% (Table 3.8).

Table 3.8

Results of Hierarchical Regression Analysis on the Problem-Solving Performance in Experiment 2

Step (Predictor)	β	R^2	ΔR^2	F for ΔR^2	p for ΔR^2
Step 1 (Proficiency)	.25**	.31	-	-	-
Step 2 (Causal Understanding)	.36**	.41	.10	10.64	.002
Step 3 (Proficiency × Causal Understanding)	.26**	.46	.05	6.21	.015

Note. ** p < .01.

To interpret this Proficiency × Causal Understanding interaction in problem-solving performance, I performed a simple slope test. Figure 3.7 shows the effect of causal understanding on the problem-solving performance as a function of readers' proficiency. It was found that better problem solving was associated with increased causal understanding when readers' proficiency was high (one *SD* above the mean of L2 reading proficiency scores), β = .58, p < .001 (see Figure 3.7). On the other hand, such a link was not found when readers' proficiency was low (one *SD* below the mean), β

= .10, p = .511. It was thus corroborated that contributions of understanding of causal relations to text learning depended on participants' L2 reading proficiency. To better understand this result, the following two sections report the results from the qualitative analyses on the causal question answers and the problem-solving responses.

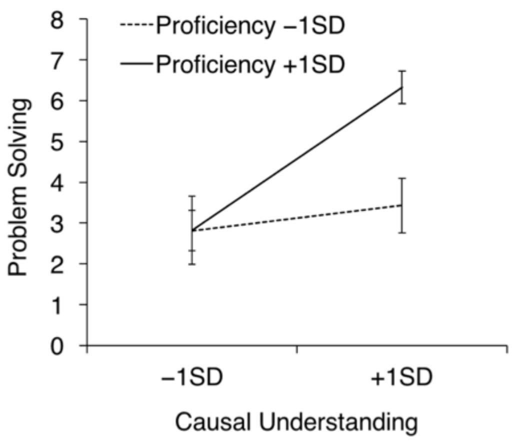

Figure 3.7. The relation between causal understanding and problem-solving performance as a function of L2 reading proficiency in Experiment 2.

3.2.3.3 Inferences in causal question answers

Table 3.9 and Figure 3.8 present the results of the analysis on three-type inferences found in causal question answers.

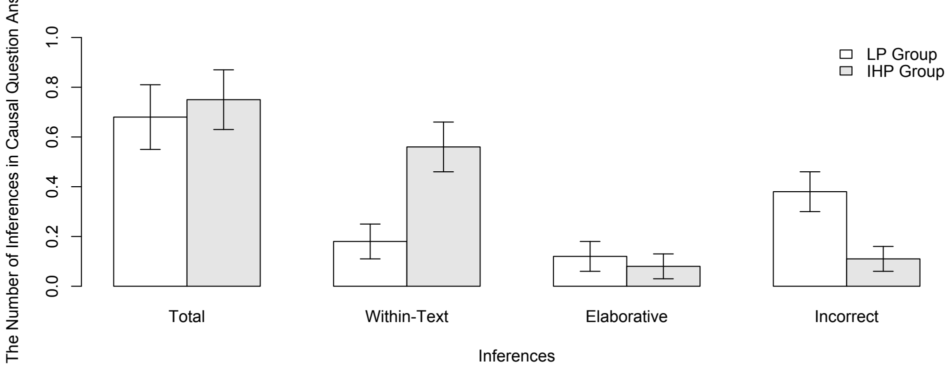

Figure 3.8. Number of inferences in causal question answers for the low-proficiency (LP) and intermediate-high-proficiency (IHP) groups ± error bars (standard errors) in Experiment 2.

Table 3.9

Numbers of Inferences in Causal-Question Answers as a Function of the Proficiency Groups and Inference Types in Experiment 2

Proficiency	*M*	95% CI	*SD*
		Total amount of inferences	
LP (*n* =34)	0.68	[0.40, 0.90]	0.77
IHP (*n* = 36)	0.75	[0.46, 0.91]	0.73
		Within-text inferences	
LP (*n* =34)	0.18	[0.04, 0.26]	0.39
IHP (*n* = 36)	0.56	[0.33, 0.70]	0.61
		Elaborative inferences	
LP (*n* =34)	0.12	[0.02, 0.23]	0.33
IHP (*n* = 36)	0.08	[−0.02, 0.14]	0.28
		Incorrect inferences	
LP (*n* =34)	0.38	[0.22, 0.53]	0.49
IHP (*n* = 36)	0.11	[0.01, 0.22]	0.32

Note. LP = low proficiency; IHP = intermediate-high proficiency; CI = confidence interval.

Due to the standard deviations being large, I used the non-parametric Mann-Whitney *U* test for statistical analyses. The total amount of inferences was not significantly different between the two proficiency groups, $U = 570$, $z = 0.54$, $p = .590$, $r = .06$. Likewise, proportions of participants who produced at least one inference were not significantly different between the groups: the intermediate-high-proficiency group at 58% (21 out of 36) and the low-proficiency group at 53% (18 out of 34), $\chi^2(1) = 0.21$, $p = .650$, $\varphi = 0.05$.

However, proficiency effects emerged within respective types of inferences. Specifically, the intermediate-high-proficiency group showed more within-text inferences than the low-proficiency group, $U = 408$, $z = 2.90$, $p = .003$, $r = .35$. Oppositely, the low-proficiency group showed more incorrect inferences than the intermediate-high-proficiency group, $U = 446$, $z = 2.63$, $p = .009$, $r = .31$. This was augmented by the fact that proportions of incorrect inferences were higher for the low-proficiency ($M = .33$, $SD = .46$) than the intermediate-high-proficiency group ($M = .08$, $SD = .25$), $U = 440$, $z = 2.70$, $p = .007$, $r = .32$. There was no significant intergroup difference for elaborative inferences, $U = 591$, $z = 0.47$, $p = .635$, $r = .06$.

3.2.3.4 Distinctive patterns of problem solving in the low-proficiency group

Scrutiny into participants' problem-solving responses revealed two distinctive patterns in the low-proficiency group. First, the low-proficiency group tended to answer events in a causally incorrect order. This was observed for 47% (16 out of 34) of the low-proficiency group and 19% (7 out of 36) of the intermediate-high-proficiency group, $\chi^2(1) = 6.04$, $p = .013$, $\varphi = 0.29$. For example, a low-proficiency-group participant answered the question in Table 1 that "Usually, the heart shrinks in space in order to reduce the body fluids. With that disease....". In the underlined part, the cause-effect order between the amount of body fluid and the heart size was reversed; in fact, the heart shrinkage is the *outcome* of the decreased body fluids, not the cause.

Second, the low-proficiency group tended to answer using information that is irrelevant to situations conveyed by the text. This was observed for 41% (14 out of 34) of the low-proficiency group and 11% (4 out of 36) of the intermediate-high-proficiency group, $\chi^2(1) = 8.27$, $p = .004$, $\varphi = 0.34$. A low-proficiency-group participant answered the question in Table 1 that "If people did not notice changes in body fluids, the amount of exercise would not differ from on the earth. So, the heart would not shrink." The relation between the amount of body fluids and the amount of exercise was not implied, let alone stated, by the text. Taken together, these patterns of incorrect problem solving suggest that low-proficiency readers failed to appropriately learn causal relations in the text as knowledge.

3.2.4 Discussion

It was confirmed that contributions of causal understanding to text learning depended on readers' L2 reading proficiency. Specifically, the results from the regression analysis showed that causal-question performance predicted problem solving when participants' L2 reading proficiency was high (one *SD* above the mean of proficiency test scores) but not when L2 reading proficiency was low (one *SD* below the mean of proficiency test scores). This is consistent with the prediction that difficulty with inferential processing would prevent low-proficiency readers from learning causal relations from the text. The combination of quantitative and qualitative analyses in the present experiment specifically provided two explanations for this finding, as follows.

The first concerns understanding of causal relations. That is, low-proficiency readers had difficulty understanding causal relations in the text. This view comes from the fact that the amount of causal relations correctly recalled in the causal question was much lower in the low- than the intermediate-high-proficiency group. It is important to note that the causal relations needed to complete the causal question were explicitly provided in the text (see Table 3.1 for the experimental text). Hence, answering these relations can be achieved by textbase comprehension. From this perspective, many of the low-proficiency readers presumably had trouble at the textbase level.

The second is that low-proficiency readers failed to construct accurate situation models. This view comes from the results concerning inferences found in causal question answers and contents of the problem-solving responses. Regarding inferences in the causal question, the qualitative analysis of the causal question answers revealed that low-proficiency readers showed much more incorrect inferences than intermediate-high-proficiency ones. The higher amount of incorrect inferences indicates that situation models constructed by low-proficiency readers were of lower accuracy, relative to intermediate-high-proficiency readers (Barry & Lazarte, 1998). Regarding contents of the problem-solving responses, low-proficiency readers tended to perform the problem solving using inappropriate causal sequences and irrelevant information, as shown by the qualitative inspection into the problem-solving responses. These patterns of the incorrect problem solving can be interpreted as suggesting that causal relations in the text were not correctly learned as long-term memory knowledge in low-proficiency readers. Consequently, those readers were inefficient at applying the causal relations beyond the text.

Combining these two accounts, it can be assumed that low-proficiency readers had trouble with processes both at textbase and situation-model levels; they struggled to understand causal relations in the text, and failed to construct accurate situation models. In general, the successful construction of situation models depends on the appropriate understanding of what is stated in the text (i.e., building of the textbase). This fact implies that low-proficiency readers' difficulty was presumably rooted in processes associated with the building of the textbase. Processes at the textbase level include understanding individual words and phrases and interrelating those ideas to form a meaningful proposition (Kintsch, 1998). At this point, it must be noted that lexical and syntactic items of the experimental passage were highly simplified. Accordingly, it is unlikely that low-proficiency readers had difficulty understanding individual words or sentences per se. Rather, their difficulty seems to lie in the process of interrelating multiple ideas to form coherent propositional idea units. This argument is consistent with prior L2 research demonstrating that less-proficient readers often fail to build relational understanding of the text, even when they can capture the meaning of text elements individually (e.g., Horiba, 2013; Hosoda, 2014). In sum, the results suggest that LP readers' difficulty with relational understanding of the text leads to their situation models having low accuracy, which then prevented understood causal relations from being learned as knowledge.

In addition to these findings, I must note several points that need careful interpretation. First, the fact that the causal question was conducted after reading leaves open the possibility that the observed inferences were made when participants answered the question, not at the time of comprehension (i.e., during reading). Indeed, some past studies reported that when required by the post-reading task, readers may make inferences that were not generated during reading (e.g., Hosoda, 2014; Noordman et al., 1992). It is possible that low-proficiency readers might be forced to rely on off-text information when answering the causal question because their understanding of causal relations was weak (as shown by their lower causal question performance). This implies that the larger amount of incorrect inferences found for low-proficiency readers might be an artifact of the methodological feature of the causal question. This study cannot rule out this possibility because participants' on-line reading processes were not directly measured. However, it must also be noted that this possibility does not mean that low-proficiency readers generated correct inferences during reading. Rather, considering that they struggled with even textbase processes, it was presumably difficult for low-proficiency readers to make appropriate inferences during reading. This is because the success in on-line inference generation usually requires the efficient processing of explicit text (Horiba, 1996; Kintsch, 1998).

Second, the number of inferences observed in this study was much lower than that found in previous research; even intermediate-high-proficiency readers demonstrated less than one inference on average. This is, at least in part, attributable to the nature of the experimental text. The text used here was short, with highly controlled linguistic complexity. The information source for inferences was accordingly reduced, possibly restricting the number of inferences made by readers. In contrast to this study, Barry and Lazarte (1998) used longer passages and elicited on average 4.9 and 8.1 inferences in recall protocols from low-knowledge and high-knowledge readers, respectively. In addition, they reported significantly more inferences ($M = 7.6$) for passages with higher syntactic complexity (defined by the number of embedded clauses per sentence) than for passages with lower syntactic complexity ($M = 5.4$). Of course, a definitive conclusion cannot be drawn because this study did not compare different versions of the text. However, it is at least possible that longer authentic texts might have elicited more inferences.

Finally, one may argue that low-proficiency readers' difficulty might be caused by the applied feature of the problem-solving text. This argument specifically holds that low-proficiency readers might struggle to reform and transfer understood ideas to novel situations when performing the problem solving. As stated before, successful problem solving requires one to apply relations or principles underlying multiple pieces of the explicit text beyond the original text. This inherently involves readers reconstructing understood text information and transferring it in such a way that the given task can appropriately be achieved. Such complex processes required in the problem-solving test might prove another difficulty for low-proficiency readers.

This argument is indirectly supported by the *transfer appropriate processing* (Morris, Bransford, & Franks, 1977). This posits that good performance can be achieved more easily when similar processes are involved between learning and retrieval than when different processes are required. The match between learning and retrieval may be smaller in the delayed problem solving than in an immediate reproductive task. The difficulty with transferring what has been understood from the text might then impede less proficient participants' performance on the problem-solving test.

The direct investigation of this account is possible by conducting an experiment where the problem-solving test is administered immediately after reading; this will discern whether the retention or the transfer of understood causal relations is responsible for less proficient readers' difficulty in text learning. The subsequent experiment pursues this point.

3.2.5 Conclusion of Experiment 2

The purpose of Experiment 2 was to explore the relation between causal understanding and learning outcomes from expository text in EFL readers. The findings of this experiment indicate that contributions of causal understanding to text learning depend on L2 reading proficiency. Causal question performance predicted problem-solving performance in readers with intermediate-high proficiency but not readers with low proficiency. The combination of quantitative and qualitative analyses suggests that low-proficiency readers had trouble with both textbase- and situation-model-level processes; consequently, they failed to learn causal relations in the text as knowledge.

Despite these findings, I must point out several limitations of this experiment. First, this experiment cannot provide direct evidence of L2 readers' inferences during reading, as participants' inference generation was assessed by an off-line task (i.e., the causal question). Direct investigation into L2 readers' online expository reading processes will provide a more grounded account of L2 readers' moment-by-moment inferential processing. This point will be addressed in later experiments in Study 2.

In addition, it is possible that the absence of a link between causal understanding and text learning in low-proficiency readers was caused by the applied nature of the problem-solving test. The difficulty associated with reconstructing and transferring what was understood from the text might hinder low-proficiency readers in using understood causal relations in the problem-solving text. If this nature of the problem solving matters, it is possible that low-proficiency readers' causal-question performance would not still contribute to problem solving, even when the problem-solving test is administered immediately after reading. I addressed this possibility in the subsequent experiment.

3.3 Experiment 3

3.3.1 Purpose and research question

Experiment 3 was designed to explore whether the applied nature of the problem-solving test is responsible for the absence of a link between causal understanding and text learning in low-proficiency readers. The findings from quantitative and qualitative analyses in Experiment 2 revealed that low-proficiency readers struggled with (a) understanding causal relations stated in the expository text (as indicated by their lower causal-question performance) and (b) constructing accurate situation models of those relations in long-term memory (as indicated by their higher number of incorrect inferences in causal question answers, and qualitatively incorrect problem-solving responses). Consequently, their understanding of causal relations did not contribute to their final learning outcomes from the text.

At the same time, the difficulty experienced by low-proficiency readers with text learning might have been caused by the nature of the problem-solving test; successful problem solving requires reforming and productively applying learned information to novel environments (e.g., Mautone & Myers, 2001). According to the transfer-appropriate processing (Morris et al., 1977), such reconstructing and transferring processes can pose additional processing difficulty due to the difference between encoding and retrieval phases.

To explore this possibility, this experiment included a problem-solving test just after participants' reading of an experimental text (hereafter referred to as the *immediate problem-solving test*). I was specifically interested in whether a Proficiency × Causal Understanding interaction (causal-question performance predicts problem-solving performance in high- but not low-proficiency readers) would appear in participants' performance on the immediate problem-solving test.

If the applied nature of the problem-solving test were responsible, the Proficiency × Causal Understanding interaction would have been expected to affect performance on the immediate problem-solving test, as the nature of the problem-solving test was the same. On the other hand, if the Proficiency × Causal Understanding interaction were not to affect performance on the immediate problem-solving test, the absence of a link between low-proficiency readers' causal understanding and learning would not have been simply attributable to the nature of the problem-solving test. Therefore, Experiment 3 addressed the following research question:

RQ3: Does the Proficiency × Causal Understanding interaction affect EFL readers' performance on the immediate problem-solving test?

3.3.2 Method

3.3.2.1 Participants

Participants were 69 Japanese undergraduate and graduate students at University of Tsukuba and National Institute of Technology, Ibaraki College. None of them participated in the previous experiments. They were majoring in humanities, education, international studies, mathematics, medicine, physics, or psychology. Seven-teen were female and 52 were male. None of them participated in Experiment 2. Their ages ranged from 18 to 22 years old (M = 18.96, SD = 1.41). They were all native speakers of Japanese and had lived in Japan for more than 18 years. They had studied English at least for six years as part of the compulsory education in Japan. None of them had experience of studying abroad for more than two weeks.

Their self-reports indicated that participants had overall English proficiency of the CEFR A2 to B2 levels, which was generally equivalent to Experiment 2's participants'

(Council of Europe, 2001; Tannenbaum & Wylie, 2008, 2013); the TOIEC listening and reading test (*M* = 547.26, *SD* = 84.00, range = 325 to 870) and the EIKEN test (Grade 4 to Grade 1: Grade 4, *n* = 0; Grade 3, *n* = 9; Grade Pre-2, *n* = 5; Grade 2, *n* = 10; Grade Pre-1, *n* = 1; Grade 1, *n* = 0). The self-reports were collected by a paper questionnaire. Fifty-eight participants reported at least one of the scores, and the other 11 participants reported none of the scores.

Participants were grouped by means of a median split (*Mdn* = 9) into low- (*n* = 34, *M* = 5.24, *SD* = 1.48, *Min*/*Max* = 2/8; the TOIEC listening and reading test: *M* = 417.01, *SD* = 60.33, range = 325 to 625) and intermediate-high-proficiency groups (*n* = 36, *M* = 12.80, *SD* = 3.70, *Min*/*Max* = 9/20; the TOIEC listening and reading test: *M* = 662.67, *SD* = 67.36, range = 520 to 870) based on their performance on the same L2 reading proficiency test as in Experiment 2 (*M* = 9.07, *SD* = 4.73, *Min*/*Max* = 3/20). An independent *t* test confirmed that the intermediate-high-proficiency group scored significantly better on the proficiency test than the LP group, $t(67) = 11.09$, $p < .01$, $d = 2.67$ (see Table 3.10).

Table 3.10

*L2 Reading Proficiency Scores of Participants in Experiment 3 (*N *= 69)*

Proficiency	*M*	95% CI	*SD*	Min	Max
LP (*n* = 34)	5.24	[4.74, 5.73]	1.48	3	8
IHP (*n* = 35)	12.80	[11.57, 14.03]	3.70	9	20

Note. LP = low proficiency; IHP = intermediate-high proficiency; CI = confidence interval; Scores on the L2 reading proficiency test range from 0 to 26.

3.3.2.2 Material

L2 reading proficiency test. The reading test from Experiment 2 was used. The reliability of the test was acceptable, with a Cronbach's α of .80.

Text. The expository text from the previous experiments was used.

Causal question. The causal question from the previous experiments was used.

Problem-solving test. The problem-solving test from Experiment 2 was used.

3.3.2.3 Procedure

The experiment's procedure was identical to Experiment 2's, except that participants completed the problem-solving test and L2 reading proficiency just after reading the expository text. The average time to complete the experiment was 90 minutes.

3.3.2.4 Scoring and data analyses

Causal question. Understanding of causal relations was assessed in the same way as in Experiments 1 and 2. Two Japanese graduate students (including I) independently scored 30% of the data, resulting in inter-rater agreement of 94%. After disagreements were resolved through discussion, I scored the remaining data.

Inferences in causal question answers were also scored as in Experiment 2. Two Japanese graduate students (including I) independently categorized 30% of the inferences into within-text, elaborative, and incorrect inferences. The inter-rater agreement was 92%. After disagreements were resolved through discussion, I categorized the remaining 70% of data.

Problem-solving test. Participants' problem-solving responses were assessed as in Experiment 2. Two Japanese graduate students (including I) scored 30% of the data.

The inter-rater agreement was 93%, with disagreements resolved through discussion. I scored the remaining data.

As in Experiment 2, two Japanese graduate students (including I), performed qualitative analysis on contents of participants' problem-solving responses to identify distinctive patterns of problem solving for the low-proficiency group.

3.3.3 Results

3.3.3.1 Descriptive statistics and intergroup differences

Table 3.11 presents the means and standard deviations of the two proficiency groups' performance on the causal question and the problem-solving test. As Figure 3.9 shows, the intermediate-high-proficiency group performed significantly better on the causal question than the LP group, $t(67) = 4.43$, $p < .001$, $d = 1.07$. Again, the majority of the low-proficiency group (85% [29 out of 34]) was found to correctly produce only two or fewer of the six target causal relations.

Similarly, the intermediate-high-proficiency group showed significantly better problem-solving performance than the low-proficiency group, $t(67) = 3.89$, $p < .001$, $d = 0.94$ (see Figure 3.10). These results are consistent with those of Experiment 2, demonstrating superior performance of the intermediate-high-proficiency group over the low-proficiency group on the two measures.

Table 3.11

Performance on the Causal Question and the Problem-Solving Test in Experiment 3

	Causal question			Problem solving		
Proficiency	*M*	95%CI	*SD*	*M*	95%CI	*SD*
LP (*n* = 34)	1.38	[1.02, 1.74]	1.07	2.65	[2.02, 3.28]	1.87
IHP (*n* = 35)	2.86	[2.32, 3.40]	1.63	4.89	[3.96, 5.82]	2.81

Note. LP = low proficiency; IHP = intermediate-high proficiency; CI = confidence interval; Scores on the causal question and the problem-solving test range 0 to 6 and 0 to 12, respectively.

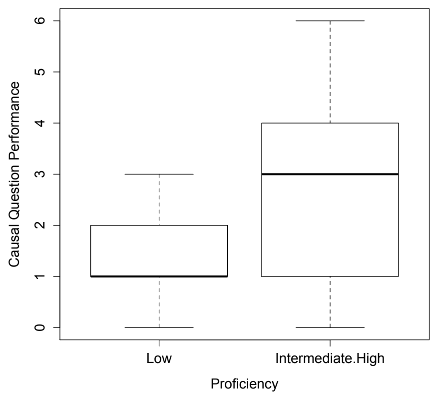

Figure 3.9. Causal-question performance in Experiment 3.

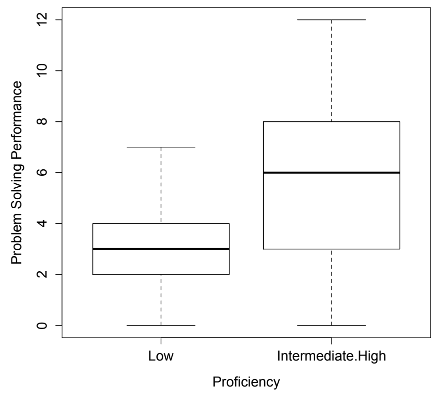

Figure 3.10. Problem-solving performance in Experiment 3.

3.3.3.2 Contributions of causal-question performance to problem solving

To examine the relation between performance on the causal question and the immediate problem-solving test, I computed correlation analyses of causal question and problem-solving performance separately for the two proficiency groups. In contrast to Experiment 2's findings, performance on the causal question and the problem-solving test were significantly correlated in the both intermediate-high- (r = 78, p < .001) and low-proficiency groups (r = 67, p < .001), as shown in Figures 3.11 and 3.12.

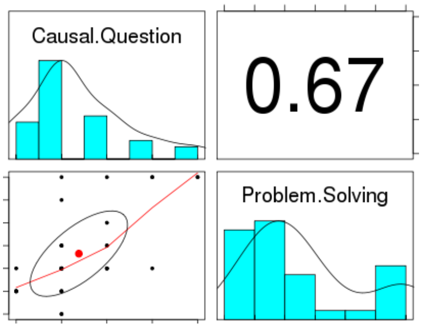

Figure 3.11. The scatter plot of the correlation between performance on the causal question and the problem-solving test in the low-proficiency group in Experiment 3.

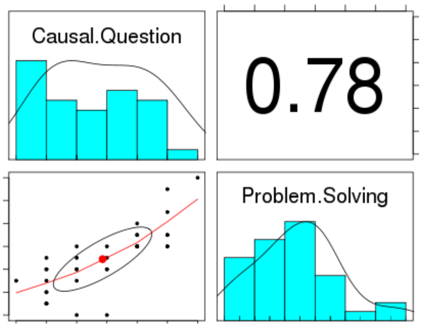

Figure 3.12. The scatter plot of the correlation between performance on the causal question and the problem-solving test in the intermediate-high-proficiency group in Experiment 3.

Next, to examine the Proficiency × Causal Understanding interaction in performance on the immediate problem-solving test, a hierarchical regression analysis was performed using immediate problem-solving performance as a dependent variable. In Steps 1 and 2, L2 reading proficiency test scores (Proficiency) and causal-question performance (Causal Understanding) were entered as predictor variables, respectively. In Step 3, I entered the interaction term of Proficiency and Causal Understanding. As in Experiment 2, this interaction term was created by multiplying L2 reading proficiency test scores by causal-question performance.

As shown in Table 3.12, the Proficiency × Causal Understanding interaction was not significant, β = .07, p = .41. Causal-question performance, on the other hand, proved highly predictive of the variance of performance on the immediate problem-solving test, β = .74, p < .001, increasing the regression model's predictive power by 43% (see Figure 3.13).

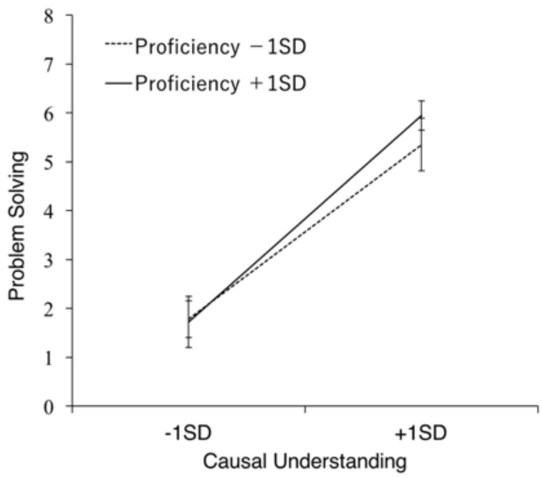

Figure 3.13. The relation between the understanding of causal relations and the problem-solving performance as a function of L2 reading proficiency in Experiment 3.

Table 3.12

Results of Hierarchical Regression Analysis on the Immediate Problem-Solving Performance in Experiment 3

Step (Predictor)	β	R^2	ΔR^2	F for ΔR^2	p for ΔR^2
Step 1 (Proficiency)	.05	.22	-	-	-
Step 2 (Causal Understanding)	.74**	.64	.43	78.68	< .001
Step 3 (Proficiency × Causal Understanding)	.07	.65	.01	0.68	.41

Note. $^{**}p < .01$.

3.3.3.3 Inferences in causal question answers

Table 3.13 and Figure 3.14 summarize the results of the analysis of the three types of inferences in causal question answers.

Table 3.13

Numbers of Inferences in Causal-Question Answers as a Function of the Proficiency Groups and Inference Types in Experiment 3

Proficiency	*M*	95% CI	*SD*
		Total amount of inferences	
LP (*n* =34)	0.71	[0.46, 0.95]	0.72
IHP (*n* = 35)	0.69	[0.46, 0.91]	0.68
		Within-text inferences	
LP (*n* =34)	0.12	[0.01, 0.23]	0.33
IHP (*n* = 35)	0.51	[0.33, 0.70]	0.56
		Elaborative inferences	
LP (*n* =34)	0.12	[0.01, 0.23]	0.33
IHP (*n* = 35)	0.06	[−0.02, 0.14]	0.24
		Incorrect inferences	
LP (*n* =34)	0.47	[0.30, 0.64]	0.51
IHP (*n* = 35)	0.11	[0.01, 0.22]	0.32

Note. LP = low proficiency; IHP = intermediate-high proficiency; CI = confidence interval.

Overall, these inference results were in line with Experiment 2's. A Mann-Whitney *U* test indicated that the total amount of inferences was not significantly different between the intermediate-high- and low-proficiency groups, $U = 570$, $z = 0.54$, $p = .590$, $r = .06$. Similarly, there was no significant group difference in proportions of participants who produced at least one inference: the intermediate-high-proficiency group at 46% (16 out of 35) and the low-proficiency group at 50% (17 out of 34), $\chi^2(1) = 0.21$, $p = .650$, $\varphi = 0.05$.

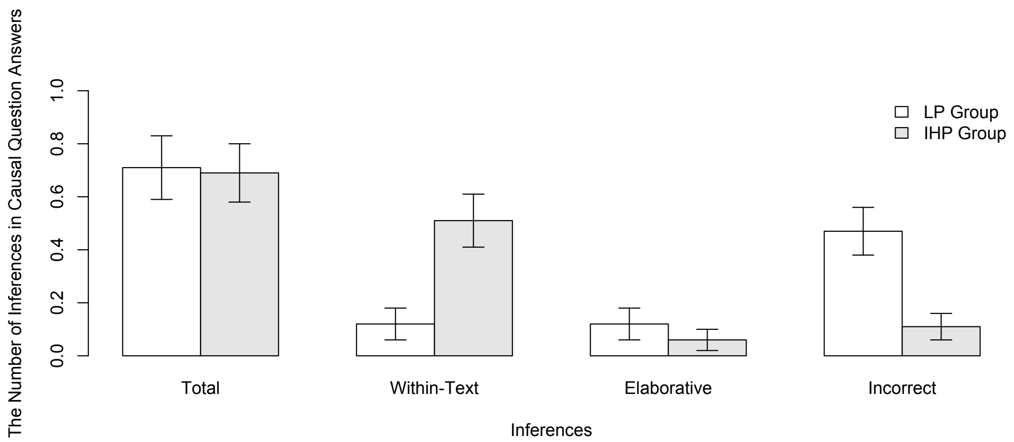

Figure 3. 14. Number of inferences in causal question answers for the low-proficiency (LP) and intermediate-high-proficiency (IHP) groups ± error bars (standard errors) in Experiment 3.

As in Experiment 2, group differences emerged in respective type of inference. The intermediate-high-proficiency group was found to make more within-text inferences than the LP group, $U = 408$, $z = 2.90$, $p = .003$, $r = .35$. On the other hand, the low-proficiency group made more incorrect inferences than the intermediate-high-proficiency group, $U = 446$, $z = 2.63$, $p = .009$, $r = .31$. Proportions of incorrect inferences were also higher for the low-proficiency group ($M = .33$, $SD = .46$) than the intermediate-high-proficiency group ($M = .08$, $SD = .25$), $U = 440$, $z = 2.70$, $p = .007$, $r = .32$. The amount of elaborative inferences was not different between the intermediate-high-proficiency and low-proficiency groups, $U = 591$, $z = 0.47$, $p = .635$, $r = .06$.

3.3.3.4 Distinctive patterns of problem solving in the low-proficiency group

In contrast to Experiment 2, only one problem-solving pattern was found to significantly characterize the low-proficiency group. The participants in the low-proficiency group responded to the problem-solving test with incorrect causal sequences more frequently than those in the intermediate-high-proficiency group. This was observed for 38% (13 out of 34) of the low-proficiency group and 14% (5 out of 36) of the intermediate-high-proficiency group, $\chi^2(1) = 5.43$, $p = .020$, $\varphi = 0.28$.

On the other hand, group differences remained marginally significant in proportions of participants who used irrelevant information in their problem-solving responses. This was observed for 24% (8 out of 34) of the low-proficiency group and 8% (3 out of 36) of the intermediate-high-proficiency group, $\chi^2(1) = 3.04$, $p = .081$, $\varphi = 0.18$.

3.3.4 Discussion

Regarding RQ3, it was found that the Proficiency × Causal Understanding interaction did not affect performance on the immediate problem-solving test. Specifically, the regression analysis revealed that causal-question performance predicted problem-solving performance, irrespective of participants' L2 reading proficiency. This finding provides evidence that the absence of a link between causal understanding and text learning in low-proficiency readers, as observed in Experiment 2, cannot be explained by the nature of the problem-solving test—Experiments 2 and 3 used the same problem-solving test, differing only in the timing of the task (i.e., Experiment 2, delayed; Experiment 3, immediate). Hence, the fact that low-proficiency readers' causal understanding did not contribute to text learning in Experiment 2 was attributable to low-proficiency readers' difficulty with (a) understanding causal relations in the text in question and (b) constructing situation models in long-term memory that properly represent the text's causal relations.

In fact, the low-proficiency group in Experiment 3 exhibited these kinds of difficulty, as indicated by two findings. First, performance on the causal question was much poorer in the low-proficiency group than in intermediate-high-proficiency group; most low-proficiency readers correctly produced only one-third or less of the key causal relations in the text. This finding indicates that low-proficiency readers struggled with building appropriate textbase of the text's causal relations. Second, low-proficiency readers generated significantly more incorrect inferences in their causal question answers than intermediate-high-proficiency readers. This finding indicates that low-proficiency readers failed to construct causally accurate situation models (Barry & Lazarte, 1998). Together, these quantitative and qualitative findings match Experiment 2's demonstrating, demonstrating that low-proficiency readers struggled with processes at both textbase and situation-model levels.

At the same time, we must note that there was no statistically significant difference between the two proficiency groups in proportions of participants who used irrelevant information in responding to the problem-solving test. In Experiment 2, low-proficiency readers showed this problem-solving pattern significantly more often than intermediate-high-proficiency readers. A possible reason for this discrepancy relates to the availability of text information. Specifically, this experiment administered the problem-solving test immediately after reading, text information may have been more available in short-term memory than it had been in Experiment 2, where the problem-solving test had been administered one week after participants' reading. As a result, there was less need for this experiment's participants to rely on off-text information than Experiment 2's participants. It is important to note that the fact that the problem-solving test requires going beyond the text does not mean that the explicit text information is unnecessary. Rather, successful problem-solving builds on readers' understanding of the explicit text (Mautone & Myers, 2001; McCrudden et al., 2009). Therefore, correctly understanding the explicit text is a necessary condition for readers to successfully complete the problem-solving test.

3.3.5 Conclusion of Experiment 3

The goal of Experiment 3 was to examine whether the applied nature of problem-solving test explains the absence of a link between causal understanding and learning from text in low-proficiency readers. To achieve this goal, I administered the problem-solving test immediately after participants read the expository text. I then examined the presence of the Proficiency × Causal Understanding interaction in participants' performance on the immediate problem-solving test.

It was shown that the Proficiency × Causal Understanding interaction did not affect immediate problem-solving performance. Instead, causal-question performance contributed to immediate problem-solving performance, both in low- and high-proficiency readers. These findings suggest that the absence of contributions of causal-question performance to problem solving in low-proficiency readers, observed in Experiment 2, is not attributable to the nature of the problem-solving test. Rather, quantitative and qualitative analyses provide the evidence that low-proficiency readers struggled to appropriately understand causal relations in the text (as indicated by their lower performance on the causal question) and construct causally accurate situation models (as indicated by their higher proportions of incorrect inferences in causal question answers and use of incorrect causal sequences in problem solving). These results complement Experiment 2’s results, demonstrating that low-proficiency readers had trouble with textbase and situation-model processes.

Summarizing the findings from Experiments 2 and 3, we can again conclude that understanding text’s causal relations leads to final learning only in readers with sufficient L2 reading proficiency. This threshold of proficiency corresponds to about the upper end of CEFR B1 to B2 levels. When readers’ proficiency is below this threshold, the difficulty with texbase- and situation-model-level processes prevents them from learning causal relations in text as knowledge.

3.4 Summary of Study 1

Study 1 explored the relation between causal understanding and learning from text in EFL readers. First, Experiment 1 examined the relation between EFL readers' understanding of causal relations (as assessed by the causal question) and their memory for expository text, by addressing RQ1 below.

RQ1: How do the quantity and the quality of text memory differ between EFL readers who can and cannot causally explain expository text well?

The results found that text memory in readers who did or did not causally explain the expository text well was different qualitatively (how interconnected the memorization of text information is) rather than quantitatively (how much text information is memorized). This finding indicates that just remembering relevant pieces of text information is not sufficient for understanding the causal relations between them. More important, causal understanding requires readers to build networked memory representations, where pieces of information are mutually interconnected.

Experiment 2 explored contributions of EFL readers' causal understanding to learning outcomes from expository text (as assessed by the problem-solving test). I considered participants' L2 reading proficiency and addressed RQ2, shown below.

RQ2: Do contributions of EFL readers' understanding of causal relations to their learning outcomes from the text differ as a function of L2 reading proficiency?

It was found that causal understanding contributed to text learning in readers with intermediate-high L2 reading proficiency, but not in low-proficiency readers (the

Proficiency × Causal Understanding interaction in learning from text). The qualitative analyses revealed that low-proficiency readers had difficulty both (a) understanding causal relations in the text and (b) building causally accurate situation models. As a result, they failed in long-term learning of the text's causal relations.

Experiment 3 explored the possibility that the observed Proficiency × Causal Understanding interaction in text learning was due to the applied nature of the problem-solving test (in which readers are asked to reconstruct and productively transfer text information to a new situation). To this end, I conducted the problem-solving test immediately after participants' text reading (the immediate problem-solving test), and addressed RQ3, shown below.

RQ3: Does the Proficiency × Causal Understanding interaction affect EFL readers' performance on the immediate problem-solving test?

The results showed that the Proficiency × Causal Understanding interaction did not affect performance on the immediate problem-solving test. This result reinforces the view that the absence of a link between causal understanding and text learning in low-proficiency readers cannot be fully explained by the nature of the problem-solving test. The qualitative analysis again proved that low-proficiency readers had problems with understanding the text's causal relations and constructing accurate situation models of those causal relations. The findings from Study 1 can be summarized in the following three statements.

(1) Causal understanding is more dependent on the quality of text memory than the quantity of it.

(2) The relation between causal understanding and learning from text depends on L2 reading proficiency

(3) Low-proficiency readers have difficulty with both textbase- and situation-model-level processes.

So far, little research has explored how EFL readers build causal understanding and learn new knowledge from expository text. For this gap, Study 1's findings provide new information about the cognitive mechanism involved in EFL readers' expository comprehension and subsequent learning.

Note that, however, Study 1 does not provide direct information about the exact status of on-line processes implemented during EFL expository reading. For example, it remains unclear how or weather EFL readers make causal bridging inferences during expository reading, and to what extent EFL readers' on-line processes reflect the causal structure of the expository text. In order to address these gaps, I designed Study 2, which directly explored on-processes involved in EFL readers' expository comprehension. I conducted three experiments investigating (a) the conditions under which EFL readers make causal bridging inferences during expository reading (Experiment 4), (b) the relation between EFL readers' on-line processes and the expository text's causal structure (Experiments 5, 6), and (c) the relation between on-line processes and off-line causal understanding (Experiment 6).

Chapter 4

Study 2: On-Line Reading Processes Involved in EFL Readers' Expository Text Comprehension

4.1 Experiment 4

4.1.1 Purpose and research questions

The purpose of Experiment 4 was to reveal the conditions under which EFL readers generate causal bridging inferences during their reading of the expository text. The overall goal of Study 2 was to explore on-line processes involved in EFL readers' causal understanding of expository text. As a starting point for this goal, this experiment was designed to examine whether and how EFL readers generate causal bridging inferences that are necessary for building situation models of causal relations. I considered the content familiarity of text and EFL readers' L2 reading proficiency as variables.

To measure on-line causal inference generation, I used a combined measure of reading times for target sentences and correct response times for inference questions (e.g., Singer et al., 1997). This paradigm was employed because it provides different and complementary perspectives on on-line inference processing (for details see Section 2.3.1.2; Singer et al., 1997; Singer & O'Connell, 2003).

Considerable research has reported that the content familiarity of text has a large influence on inferential processing (e.g., McNamara et al., 1996). Specifically, when text content is familiar to readers, a rich amount of text-relevant information is likely to be available from their knowledge base, facilitating knowledge-driven processing (Kintsch, 1998). Such processing subsequently enables readers to smoothly incorporate text information into their existing knowledge base. Thus, the extent to which readers make use of prior knowledge to make inferences largely depends on the familiarity of text content.

At the same time, researchers point out that readers' knowledge is not always applicable to moment-by-moment on-line processes with the success determined by one's reading proficiency levels. McNamara and O'Reilly (2009) proposed that less proficient readers often fail to activate relevant knowledge because knowing when it is necessary to use such knowledge is usually difficult for them. This situation seems particularly likely for EFL readers, who often have trouble interpreting explicitly stated information in text. In fact, Morishima (2013) showed that Japanese EFL university students do not exhibit knowledge-based activation of information, and are hence are less sensitive than L1 readers to inconsistencies in the texts.

Based on these exiting findings, it is likely that the content familiarity of text and readers' L2 reading proficiency can interactively affect causal inference generation in such a way that only readers with sufficient proficiency benefit from high-familiarity text. The following RQ was addressed in the experiment.

RQ4: Under what conditions do EFL readers make causal bridging inferences during reading expository texts, when considering the content familiarity of the text and readers' L2 reading proficiency?

4.1.2 Material construction

Before the experiment, I constructed materials that have two familiarity conditions: *high-familiarity* texts, with content familiar to readers; and *low-familiarity* texts, with unfamiliar content. Readers are expected to make more causal inferences during the reading of high-familiarity texts than low-familiarity texts (Kintsch, 1998;

McNamara & O'Reilly, 2009). Conversely, it would be difficult to make inferences from low-familiarity texts because relevant knowledge available for inference processing is quite limited (e.g., O'Reilly & McNamara, 2007; McNamara et al., 1996).

To establish the two familiarity conditions, experimental passages were constructed through discussions and a pilot study. The construction of materials followed a procedure specified by Singer, et al. (1997). Specifically, I took the following six-step procedure. First, I constructed 45 *because* target sentences by selecting entries from *McGraw-Hill Encyclopedia of Science and Technology* (McGraw-Hill, 1992) that could be expressed using *because* (e.g., "It is difficult to set a fire in high mountains because there is less oxygen"). Second, surrounding contexts were created from related descriptions. Third, three Japanese graduate students and I separately identified candidate inference ideas in Japanese (e.g., 火を起こすには酸素が必要だ [Oxygen is necessary to make fires.]) that are necessary for linking causal relations in target sentences. Fourth, we had a discussion, in which each inference idea was segregated into a phrase unit in Japanese (i.e., bunsetsu; e.g., 火を/起こすには/酸素が/必要だ/). Fifth, for each inference idea, we selected the phrases that had been identified by at least three of the four raters, and these were used in a pilot study (see the following section). Sixth and finally, low-frequency words with level 5 or more according to the JACET 8000 list (Ishikawa et al., 2003) were substituted with plainer words. Several technical terms that could not be rephrased using simpler words were explained in Japanese in parentheses (e.g., *Spinach* (ほうれんそう); see Appendix B for an example); however, no target sentences in the experimental passages contained such words. Finally, a native English speaker checked the materials for naturalness.

4.1.3 Pilot study

Purpose. To determine passages that suited the experiment's purposes, a pilot study was conducted with a total of 65 expository passages. The purpose of the pilot study was twofold: (a) to identify passages satisfying the two familiarity conditions (i.e., low and high familiarity) and (b) to confirm the viability of inference generation.

Procedure. Forty separate Japanese undergraduate and graduate students at University of Tsukuba read 65 passages with inference ideas removed. The passages were presented on a PC screen one sentence at a time, and participants processed each sentence in a self-paced manner. After reading each passage, participants were asked to rate the content familiarity of that text read on a 7-point scale (1 = *totally unfamiliar*, 7 = *very familiar*; i.e., a *familiarity rating*). Participants submitted their ratings by pressing buttons numbered 1 to 7 on a response pad (Cedrus, CA, USA).

After reading all of the passages, participants were provided with a booklet containing 65 inference ideas. Participants were then asked to answer "yes" when they thought each inference idea had been stated in the texts they had just read, and "no" otherwise (a *recognition task*; Muramoto, 2000).

Scoring, data analysis, and results. Mean familiarity ratings were calculated for each passage to determine experimental passages. Those with mean ratings of 5 or more were classified as high familiarity, and those with mean ratings of 3 or less were classified low familiarity; 14 and 15 passages met the criteria for high- and low-familiarity conditions, respectively. Finally, to ensure that implicit ideas were understandable for EFL readers through inference generation, 12 high-familiarity and 12 low-familiarity passages were selected. These passages contained inference ideas that most participants judged as having appeared in the texts in the recognition task; presumably, those inference ideas were falsely recognized because participants had inferred and encoded them. Twenty-four filler texts were added, resulting in a total of 48 expository passages to be used in the main experiment.

4.1.4 Method

4.1.4.1 Participants

Participants in Experiment 4 were 55 Japanese undergraduate and graduate students at University of Tsukuba. They were majoring in education, humanities, literature, psychology, international studies, engineering, or social studies. None of them participated in the previous experiments. Twenty-one of them were female, and the other 34 were male. Their ages range 18 to 24 years (*M* = 20.30, *SD* = 1.67). All of them had studied English for more than six years. Data from five participants who experienced trouble with experimental tools were removed.

According to their self-reports, participants' overall English proficiency was estimated to the CEFR A2 to B2 levels (Council of Europe, 2001; Tannenbaum & Wylie, 2008, 2013); the TOIEC listening and reading test (*M* = 667.26, *SD* = 103.12, range = 445 to 950) and the EIKEN test (Grade 4 to Grade 1: Grade 4, *n* = 1; Grade 3, *n* = 3; Grade Pre-2, *n* = 6; Grade 2, *n* = 15; Grade Pre-1, *n* = 5; Grade 1, *n* = 2). The self-reports were collected by a paper questionnaire. Fifty participants reported at least one of the scores, and the other five participants reported none of the scores. It must be noted that, although in the same CEFR level range, Study 2's participants seemed to have generally higher proficiency than Study 1's participants, according to their proficiency test scores (see Table 3.5).

The 50 participants were grouped by means of a median split (*Mdn* = 14) into two proficiency groups, based on their scores on the same L2 proficiency test as in Experiments 2 and 3 (*M* = 14.02, *SD* = 4.43, *Min*/*Max* = 4/24). Participants who scored less than the median formed the *intermediate-low-proficiency group* (*n* = 25, *M* = 9.48, *SD* = 2.86, *Min*/*Max* = 4/13). According to their self-reports (the TOEIC listening and reading test: *M* = 542.46, *SD* = 59.25, range = 445 to 630), this group was estimated to be placed in the lower end of the CEFR B1 level, or about in the middle between the low-proficiency and intermediate-high-proficiency groups in Experiments 2 and 3 (Tannenbaum & Wylie, 2008, 2013). On the other hand, participants who scored the median or higher on the proficiency test formed the *high-proficiency group (n* = 25, *M* = 17.52, *SD* = 3.38, *Min*/*Max* = 14/26). This group was estimated to be placed in the CEFR B2 level, according to their self-reports (the TOEIC listening and reading test: *M* = 784.10, *SD* = 68.02; range = 630 to 950; Tannenbaum & Wylie, 2008, 2013), and had higher proficiency than the intermediate-high-proficiency group in Experiments 2 and 3. An independent *t* test found significant score differences between the groups $t(48) = -9.82$, $p < .001$, $d = 2.25$ (see Table 4.1).

Table 4.1

*L2 Reading Proficiency Scores of Participants in Experiment 4 (*N *= 50)*

Proficiency	*M*	95% CI	*SD*	Min	Max
ILP (*n* = 25)	9.48	[8.36, 10.60]	2.86	4	13
HP (*n* = 25)	17.52	[16.19, 18.85]	3.38	14	26

Note. ILP = intermediate-low proficiency; HP = high proficiency; CI = confidence interval; Scores on the L2 reading proficiency test range from 0 to 26.

4.1.4.2 Materials

L2 reading proficiency test. The same L2 reading proficiency test was used as in Experiments 2 and 3.

Texts. A total of 48 passages were used, of which 24 were experimental passages selected through the pilot study (see Sections 4.1.3). The half of the experimental passages were high-familiarity passages, and the other half were low-familiarity passages. The other 24 were fillers. An example experimental passage (high-familiarity condition) was shown in Table 4.2. All the experimental passages are provided in Appendix B.

Table 4.2

An Example Experimental Passage, Inference, and Detail Questions

- The presentation order in the explicit condition：(a)→(b)→(c)→(d)→(e)→(f)
- The presentation order in the implicit condition：(a) → (c)→(d)→(e)→(f)

(a) Skiers of different abilities need different equipment from compact to large.

(b) It is easier to turn on the shorter compact skis. (direct sentence)

(c) Beginners are frequently advised to use compact skis because they usually have difficulty in changing directions. (target sentence)

(d) Once they can control their movement, they can quickly advance in skill.

(e) コンパクトなスキーは曲がりやすい (inference question)

(f) スキーの初心者は上達に時間がかかる (detail question)

Note. Reading and response times for the underlined sentence and question were recorded.

As in Table 4.2, each passage contained a target sentence, in which *because* signaled a causal relation. In line with Singer et al.'s (1997) study, there was an implicit and an explicit condition. In the explicit condition, the target sentence was preceded by a *direct sentence* that directly referred to the inference idea that causally bridged the two clauses of the target sentence. In the implicit condition, the direct sentence was excluded so that participants in this condition had to infer the missing idea to achieve a causally coherent understanding of the passage.

Each experimental passage was followed by two verification questions: an inference and a detail. The inference question probed the inference to be derived from the target sentence. The detail question queried a specific detail of the passage to encourage participants' careful reading. The correct answer was "yes" for half of the detail questions, and "no" for the other half. For filler passages, the inference question was replaced with another detail question, for which the correct answer was always "no" to prevent readers from thoughtlessly answering "yes" to the first questions. These questions were written in Japanese so that L2 decoding ability, which was out of the study's scope, would not affect response times.

Three material sets were created with the 48 passages. Within a set, 24 experimental passages were evenly assigned to the two explicitness conditions (i.e., explicit, implicit). Filler passages consistently appeared in the explicit condition. The assignment of each text to the two explicitness conditions was counterbalanced across sets.

Familiarity ratings. To confirm the content familiarity of the text, familiarity ratings, shown in Figure 4.1, were created for the 24 experimental passages. The 7-point scale from the pilot (1 = *totally unfamiliar*, 7 = *very familiar*) was used. The familiarity ratings were provided on a PC screen.

英文の内容について、どれだけ馴染みがあったかを7段階で評価してください。

1	2	3	4	5	6	7
馴染みがない	馴染みが少ない	あまり馴染みがない	どちらともいえない	やや馴染みがある	馴染みがある	とても馴染みがある

Figure 4.1. An example familiarity rating.

4.1.4.3 Procedure

Participants were tested in a group of one to four people. First, the experimenter explained the purpose and general procedure of the experiment, and the informed consent was obtained. A reading task was then administered.

In the reading task, a "*ready*?" signal was displayed at the center of the PC screen. After participants pressed a "yes" button on the Response-Pad RB-730 (Cedrus, CA, USA), the first sentence appeared. Participants were instructed to read each sentence for understanding in a self-paced manner. Reading times for target sentences were recorded with Super Lab 4.5 (Cedrus, CA, USA). After one passage was processed, the screen went blank for 2,500 milliseconds (ms) and a 500-ms fixation (***) followed. Then, the first verification question appeared. Participants were asked to verify the question as quickly and accurately as possible by pressing the "yes" or "no" button. Response times and accuracy were recorded. Following another blank (2,500 ms) and a fixation (500 ms), the second verification question appeared. After answering this second question, another fixation (###) was presented for 500 ms, and then participants were asked to rate the familiarity of the text on a 7-point scale. After a 1000-ms blank, another "*ready*?" signal was presented. This sequence was repeated for 48 texts that were presented in a random order. Participants took a five-minute break halfway through the reading task.

After the reading task was finished, participants took the L2 reading proficiency test for 30 minutes. The average time to complete the experiment was 80 minutes.

4.1.4.4 Scoring and data analyses

The joint pattern of response times for inference questions and reading times for target sentences was used to index on-line causal bridging inference generation (Noordman et al., 1992; Singer et al., 1997). If readers generate inferences during reading, reading times for target sentences should be significantly longer in the implicit condition than in the explicit condition, due to the mental effort needed to infer the missing idea. On the other hand, response times for inference questions should be almost the same across the implicit and explicit conditions because the question queried inference information.

In the analysis, target reading and inference response times were averaged and divided by the number of syllables to account for differences in sentence and question length. Next, data were removed from those cells that received mean familiarity ratings of 5 or more if they were originally intended as the low-familiarity condition, and 3 or less if they were intended as the high-familiarity condition. Then, reading and response times less than 100 ms or three standard deviations away from the mean that were for a given participant were removed. These omissions resulted in the removal of 6.25 % of the dataset.

4.1.5 Results

4.1.5.1 Familiarity ratings

It is necessary to confirm that the familiarity of the experimental texts followed an intended direction. Table 4.3 shows the results. Overall, mean text familiarity ratings for high- and low-familiarity texts were 5.21 and 2.57, respectively. To examine rating differences between the two proficiency groups and between the two familiarity conditions, a 2 (Proficiency: high, intermediate-low) × 2 (Familiarity: high, low) two-way ANOVA was performed familiarity ratings. The results revealed a significant main effect of the familiarity, $F(1, 22) = 755.66$, $p < .001$, $\eta_p^2 = .97$, indicating that content familiarity was significantly higher for high-familiarity than for low-familiarity texts. No other effects reached statistical significance (all ps > .050).

It was therefore confirmed that participants perceived the familiarity of texts in the two familiarity conditions in an expected way; high-familiarity texts were more familiar than low-familiarity texts, and vise verse. In addition, this was the case for both the high- and intermediate-low-proficiency groups (i.e., no effects involving the proficiency factor were significant).

Table 4.3

Familiarity Ratings as a Function of the Familiarity of Text Content and the Proficiency Groups

	High familiarity			Low familiarity		
Proficiency	*M*	95% CI	SD	*M*	95% CI	*SD*
ILP (*n* =25)	5.17	[5.03, 5.31]	0.24	2.65	[2.51, 2.77]	0.23
HP (*n* = 25)	5.24	[5.05, 5.35]	0.27	2.48	[2.31, 2.66]	0.31

Note. ILP = intermediate-low proficiency; HP = high proficiency; CI = confidence interval; Familiarity ratings range from 1 (*totally unfamiliar*) to 7 (*very familiar*).

4.1.5.2 Measures of on-line generation of causal bridging inferences

Target reading times. Table 4.4 shows the means, standard deviations, and 95% CIs for reading times for target sentences. To examine whether participants made causal bridging inferences during reading, target reading times were submitted to a three-way 2 (Proficiency: high, intermediate-low) × 2 (Familiarity: high, low) × 2 (Explicitness: implicit, explicit) ANOVA.

The results showed a significant main effect of the proficiency, $F(1, 48) = 10.04$, $p = .003$, $\eta_p^2 = .17$, indicating that the high-proficiency group read target sentences faster than the intermediate-low-proficiency group. There was also a significant main effect of the explicitness, $F(1, 48) = 18.00$, $p < .001$, $\eta_p^2 = .27$. However, this explicitness effect was modified by a significant Familiarity × Explicitness interaction, $F(1, 199) = 12.12$, $p = .001$, $\eta_p^2 = .20$.

Table 4.4

Reading Times for Target Sentences (in milliseconds) as a Function of the Familiarity of Text Content and the Proficiency Groups

	High familiarity			Low familiarity		
Explicitness	*M*	95% CI	*SD*	*M*	95% CI	*SD*
			ILP group (*n* = 25)			
Explicit	384.56	[330.13, 438.99]	138.85	471.70	[443.90, 526.20]	139.02
Implicit	495.94	[438.58, 553.31]	146.35	492.38	[449.82, 534.94]	108.56
			HP group (*n* = 25)			
Explicit	339.59	[305.85, 373.33]	86.07	368.22	[331.75, 404.68]	93.03
Implicit	416.63	[370.43, 462.84]	117.87	401.93	[359.04, 444.82]	109.42

Note. ILP = intermediate-low proficiency; HP = high proficiency; CI = confidence interval.

I performed follow-up tests on this Familiarity × Explicitness interaction and found that target reading times were longer in the implicit than in the explicit condition for high-familiarity texts, $F(1, 48) = 28.05$, $p < .001$, $\eta_p^2 = .37$ (see Figures 4.2 and 4.3). By contrast, target reading times did not significantly differ between the two explicitness conditions for low-familiarity texts, $F(1, 48) = 2.66$, $p = .110$, $\eta_p^2 = .05$. It was therefore found that target reading times supported causal bridging inference generation for high-familiarity texts (i.e., implicit > explicit) but not for low-familiarity texts (i.e., implicit ≒ explicit). In addition, these results held between the two proficiency groups, as revealed by the absence of a Proficiency × Explicitness interaction, $F(1, 199) = 0.14$, $p = .771$, $\eta_p^2 = .01$.

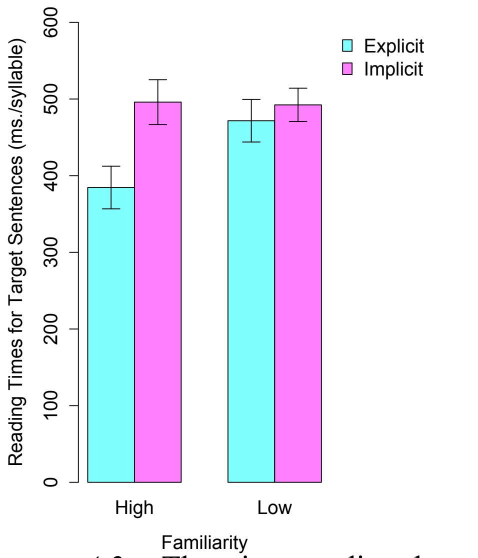

Figure 4.2. The intermediate-low-proficiency group's target reading times ± error bars (standard errors).

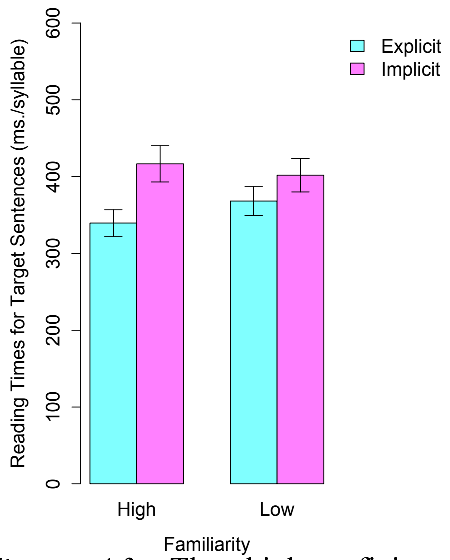

Figure 4.3. The high-proficiency group's target reading times ± error bars (standard errors).

Correct inference response times and inference error rates. Table 4.5 shows the means, standard deviations, and 95% CIs for correct response times for inference questions. Inference response times were submitted to the three-way 2 (Proficiency: high, intermediate-low) × 2 (Familiarity: high, low) × 2 (Explicitness: explicit, implicit) ANOVA. The results showed a main effect of the proficiency factor, $F(1, 48) = 6.11$, $p = .017$, $\eta_p^2 = .11$, indicating that the high-proficiency group responded to the inference question faster than the intermediate-low-proficiency group. I additionally found main effects of the familiarity, $F(1, 48) = 26.51$, $p < .001$, $\eta_p^2 = .36$, and the explicitness, $F(1, 48) = 20.46$, $p < .001$, $\eta_p^2 = .30$. However, these were qualified by a significant

Proficiency × Familiarity × Explicitness interaction, $F(1, 199) = 4.62$, $p = .037$, $\eta_p^2 = .09$. To interpret this interaction, I ran follow-up tests focusing on an aspect of findings related to the research question, that is, the Familiarity × Explicitness interaction within each proficiency group.

Table 4.5

Correct Response Times for Inference Questions (in milliseconds) as a Function of the Familiarity of Text Content and the Proficiency Groups

	High familiarity			Low familiarity		
Explicitness	*M*	95% CI	*SD*	*M*	95% CI	*SD*
			ILP group (*n* = 25)			
Explicit	111.62	[102.75, 120.49]	22.62	143.64	[131.68, 155.59]	30.50
Implicit	133.62	[120.30, 146.50]	34.00	156.32	[144.18, 168.47]	30.98
			HP group (*n* = 25)			
Explicit	106.68	[95.86, 117.50]	27.61	116.22	[101.17, 131.26]	38.38
Implicit	111.61	[102.22, 121.00]	23.95	145.76	[123.41, 168.11]	57.01

Note. ILP = intermediate-low proficiency; HP = high proficiency; CI = confidence interval.

For the intermediate-low-proficiency group, the Familiarity × Explicitness interaction was not significant, $F(1, 24) = 0.93$, $p = .345$, $\eta_p^2 = .04$ (see Figure 4.4). In contrast, the main effect of the explicitness factor was significant, $F(1, 24) = 14.24$, $p = .001$, $\eta_p^2 = .37$, indicating that the intermediate-low-proficiency group took significantly longer to verify inference questions in the implicit condition than in the explicit condition regardless of the familiarity condition. This is incompatible with the on-line inference criterion, which suggests that response times should be comparable across the explicitness conditions. Hence, these results suggest that intermediate-low-proficiency readers failed to make causal inferences during reading expository texts.

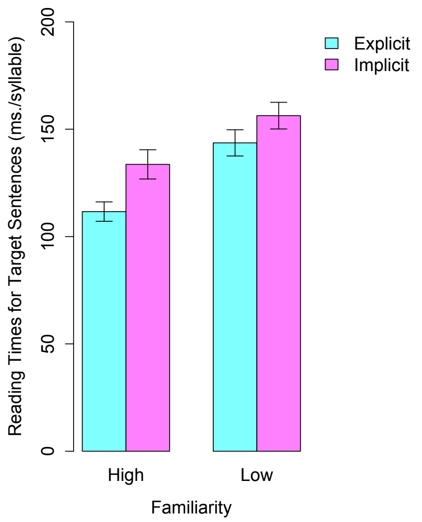

Figure 4.4. The intermediate-low-proficiency group's inference response times ± error bars (standard errors).

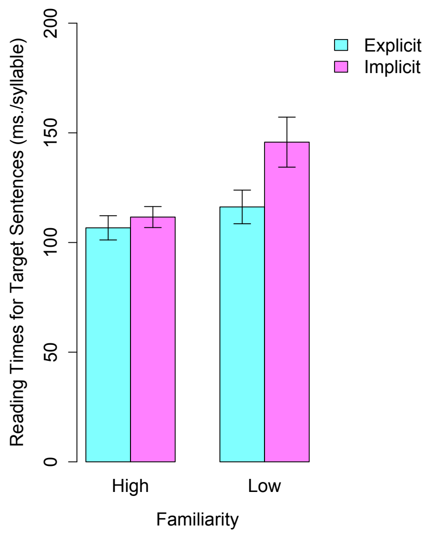

Figure 4.5. The high-proficiency group's inference response times ± error bars (standard errors).

Next, for the high-proficiency group, the Familiarity × Explicitness interaction was significant, $F(1, 24) = 4.44$, $p = .046$, $\eta_p^2 = .16$ (see Figure 4.5). Follow-up tests revealed that inference response times did not significantly differ between the explicitness conditions for high-familiarity texts, $F(1, 24) = 1.02$, $p = .322$, $\eta_p^2 = .04$, but were significantly delayed in the implicit condition for low-familiarity texts, $F(1, 24) = 6.77$, $p = .016$, $\eta_p^2 = .22$. Thus, inference response times in the high-proficiency group support on-line causal inference generation for high-familiarity texts (i.e., implicit ≒ explicit), but not for low-familiarity texts (i.e., implicit > explicit).

Finally, I analyzed error rates for inference questions to examine the accuracy of inference generation. Table 4.6 shows mean inference error rates.

Table 4.6

Error Rates for Inference Questions as a Function of the Familiarity of Text Content and the Proficiency Groups

	High familiarity			Low familiarity		
Explicitness	*M*	95% CI	*SD*	*M*	95% CI	*SD*
			ILP group (n = 25)			
Explicit	.08	[.01, .14]	.16	.11	[.05, .17]	.16
Implicit	.20	[.10, .30]	.25	.28	[.19, .37]	.23
			HP group (n = 25)			
Explicit	.06	[.02, .10]	.11	.10	[.04, .16]	.16
Implicit	.08	[.03, .13]	.12	.15	[.06, .24]	.24

Note. ILP = intermediate-low proficiency; HP = high proficiency; CI = confidence interval.

The three-way 2 (Proficiency: high, intermediate-low) × 2 (Familiarity: high, low) × 2 (Explicitness: explicit, implicit) ANOVA on inference error rates revealed significant main effects of the familiarity, $F(1, 48) = 5.14$, $p = .028$, $\eta_p^2 = .10$, indicating that error rates were higher for low-familiarity than for high-familiarity texts. There were also main effects of the proficiency, $F(1, 48) = 4.29$, $p = .044$, $\eta_p^2 = .08$, and the explicitness factors, $F(1, 48) = 14.07$, $p < .001$, $\eta_p^2 = .23$. Yet, these two main effects were modified by a significant Proficiency × Explicitness interaction, $F(1, 48) = 5.26$, $p = .026$, $\eta_p^2 = .10$.

Follow-up tests on this Proficiency × Explicitness interaction showed that the intermediate-low-proficiency group made more errors than the high-proficiency group for high-familiarity texts, $F(1, 48) = 6.84$, $p = .012$, $\eta_p^2 = .13$. On the other hand, error rates were not significantly different between the two proficiency groups for low-familiarity texts, $F(1, 48) = 0.19$, $p = .661$, $\eta_p^2 = .01$.

To be more specific, as shown in Table 4.6, this Proficiency × Explicitness interaction seems to be guided by the intermediate-low-proficiency group's higher error rates for high-familiarity texts in the implicit condition ($M = .20$, $SD = .25$), as relative to the high-proficiency group ($M = .08$, $SD = .12$). This suggests that the accuracy of inferential processing was lower in the intermediate-low- than in the high-proficiency group. Finally, inference error rates were significantly higher in the implicit condition than in the explicit condition for the intermediate-low-proficiency group, $F(1, 24) = 12.74$, $p = .002$, $\eta_p^2 = .35$, but were not significantly different between the two explicitness conditions for the high-proficiency group, $F(1, 24) = 1.88$, $p = .183$, $\eta_p^2 = .07$.

4.1.6 Discussion

In what follows, I interpret the present experiment's results to provide discussion regarding RQ4 (Under what conditions do EFL readers make causal bridging inferences during reading expository texts, when considering the content familiarity of the text and readers' L2 reading proficiency?). First, each of the two proficiency groups are discussed. After that, I discuss the regarding low-familiarity text results.

4.1.6.1 Intermediate-low-proficiency group's failure to make causal bridging inferences

For the intermediate-low-proficiency group, causal bridging inference generation was not supported in either familiarity condition, as indicated by the fact that they took significantly longer times to verify inference questions in the implicit condition than in the explicit condition. Because intermediate-low-proficiency readers failed to make causal bridging inferences during reading, they needed more time to answer inference questions when those inference ideas were not stated in texts than when those ideas were stated (e.g., Noordman et al., 1992; Singer et al., 1997).

At the same time, it is also noteworthy that intermediate-low-proficiency readers showed significantly longer target reading times in the implicit condition than in the explicit condition for high-familiarity texts, consistent with on-line inference criterion. The increased reading times in the implicit condition imply the possibility that intermediate-low-proficiency readers of high-familiarity texts recognized causal gaps in target sentences in the implicit condition. It is widely recognized that reading times are longer when readers notice textual inconsistencies or coherence breaks during reading (e.g., Morishima, 2013; O'Brien et al., 1998), because such inconsistencies introduce difficulty in incorporating the incoming information into existing mental representations. This account implies that intermediate-low-proficiency readers might be aware of some inconsistencies associated with target sentences in implicit-high-familiarity texts.

However, it is still unlikely that this group generated appropriate inferences during reading. This reasoning is because intermediate-low-proficiency readers' inference error rates were quite higher in the implicit than in the explicit condition even for high-familiarity texts. Their high inference error rates, coupled with their delayed inference response times in the implicit condition, suggest that appropriate inference ideas were not available when they answered inference questions. In other words, when reading high-familiarity texts, intermediate-low-proficiency readers arrived at the detection of causal gaps (as reflected in their increased target reading times in the implicit condition), but stopped there; they did not make appropriate causal bridging inferences to fill in those gaps. Consequently, inference response times and error rates were increased in the implicit than in the explicit condition.

The absence of inference generation in the intermediate-low-proficiency group suggests that their on-line processes confined to a textbase or explicit text level. This conclusion means that intermediate-low-proficiency set standards of coherence on explicitly stated text ideas, not on implicit relations between them. Noteworthy regarding this argument is the fact that intermediate-low-proficiency readers failed to make causal bridging inferences even from texts whose content was relatively familiar (i.e., high-familiarity texts). Specifically, apart from high content familiarity, the present experimental texts were short and linguistically controlled. It is therefore unlikely that lower-level linguistic processes posed high cognitive demands. This means that the lack of cognitive resources alone cannot fully explain the intermediate-low-proficiency group's results. In addition to the limitation of cognitive resources, the intermediate-high-proficiency group might place their standards of coherence on the interpretation of explicit text information. They were consequently satisfied when they captured explicitly stated information, thus did not attempt to infer implicit information even from familiar simplified texts. Previous research indeed indicated that less-skilled readers often have low standards of coherence on semantic similarities (Wittwer & Ihme, 2014).

The intermediate-low-proficiency group's results for high-familiarity texts additionally means that they failed to make use of the familiarity of texts to facilitate their inferential processes. This observation is in line with the view from L1 reading literature that maintains that readers needs to have a certain level of reading proficiency before they can benefit from the familiarity of text content (e.g., McNamara & O'Reilly, 2009). Because intermediate-low-proficiency readers were forced to focus on textbase-level processes (e.g., understanding individual explicit text ideas, integrating individual ideas to forma a proposition), they could not use high familiarity to facilitate situation-model-level processes, including inference generation.

Theoretically speaking, the existing discourse models hold that successful situation model processes, including inference generation, build on the completion of textbase processes (e.g., Graesser et al., 1994; Kintsch, 1998). From this perspective, the difficulty with the intermediate-high-proficiency group's causal inference generation might be attributable to other lower cognitive processes involved in building textbase. For example, understanding causal relations between clauses inherently involves referential resolution between those clauses. It is possible that the intermediate-high-proficiency group struggled with such lower processes, with the result being that their causal inference generation was impaired.

It is also possible that the intermediate-low-proficiency group could not effectively use the causal connective *because* to guide their inferential processing. According to Noordman and Vonk (1997), *because* has three different functions that guide readers' causal interpretations. The first is the *segmentation* function, by which *because* signals grammatical structure of the clauses (e.g., the clause beginning with *because* is a subordinate clause) to readers. The second is the *integration* function; *because* specifies that the two current clauses are integrated by means of a causal relation. The third is the *inference* function; *because* instructs readers to infer linking ideas that underlie the candidate causal relation, that is, to make causal inferences. Within this framework, it is assumed that the intermediate-low-proficiency group reached the segmentation and integration stages, but not the inference stage. This argument fits well with the result discussed above in that only target reading times supported inference generation in the intermediate-low-proficiency group; they noticed causal gaps in causal relations signaled by *because*, leading to delayed target reading times in the implicit condition. However, their processing went no further than detecting signaled causal relations; consequently, they did not utilize the integration or inference functions of *because*.

4.1.6.2 High-proficiency group's causal bridging inference generation for high-familiarity texts

The joint pattern of target reading times and inference response times revealed that the high-proficiency group made causal bridging inferences during reading high-familiarity texts. Specifically, reading times for target sentences for those texts were significantly longer in the implicit condition than in the explicit condition. On the other hand, their correct response times for inference questions were similar across the two explicitness conditions. This pair of results collectively indicate that participants made causal bridging inferences during reading on the condition that the content familiarity of the text and readers' L2 reading proficiency were both high.

The result that high-proficiency readers successfully generated causal bridging inferences from high-familiarity texts is consistent with past L2 studies on L2 inference generation (e.g., Shimizu, 2015; Yoshida, 2003), which reported that proficient L2 readers make inferences during narrative reading more than less proficient readers. However, most past studies on EFL readers' inference generation have used narratives that deal with relatively familiar topics to readers, such as everyday experiences or characters' goal-action relations (e.g., Horiba, 1996; 2000). On the other hand, this study used expository texts that contained information on scientific phenomena and showed the conditions under which EFL readers make causal inferences during reading. In this respect, this study offered a new understanding of EFL readers' inferential processing in expository comprehension.

In theoretical terms, high-proficiency readers' causal inference generation supports the assumption of readers' drive for causal reasoning proposed in existing models of discourse processing (e.g., the causal network model: Suh & Trabasso, 1993; Trabasso et al., 1989; the constructionist view: Graesser et al., 1994). According to this assumption, readers routinely try to explain why events or phenomena occur in a text or why the author mentioned them during text comprehension. This view implies that readers make inferences to explain the described text situation, when the explanation provided by the explicit text is insufficient. This situation is exactly what participants were supposed to encounter in target sentences in the implicit condition. For example, in the target sentence "Beginners are frequently advised to use compact skis because they usually have difficulty in changing directions.", the explicit text does not provide sufficient explanation for why the subordinate clause led to the main clause; that is, why compact skis are good for those who cannot turn well. The causal reasoning assumption here proposes that readers of this target sentence infer an implicit idea (*It is easier to turn on compact skis*.) that explains the causal relation conveyed in that target sentence. High-proficiency readers' on-line inference generation thus indicates that EFL readers with sufficient L2 reading proficiency do not settle for the understanding of the explicit text, but search for information that provides sufficient explanation for what they are reading. This view clearly shows that high-proficiency readers have standards of coherence on the understanding of causal relations between events in texts, not on the mere interpretation of individual sentences. This observation matches findings from the recent L1 research showing that skilled L1 readers base their standards of coherence on causality among multiple sentences (Wittwer & Ihme, 2014).

The high-proficiency group's inference generation from high-familiarity texts also implies that they successfully benefited from high familiarity of text content to guide their on-line inferential processing. This finding stands in stark contrast to that for the intermediate-low-proficiency group, who failed to make causal bridging inferences even from high-familiarity texts. In fact, high-proficiency readers showed quite higher accuracy in their inference generation from high-familiarity texts than intermediate-low-proficiency readers. This argument comes from the fact that high-proficiency readers' inference error rates for implicit-high-familiarity texts were lower (M = .08, SD = .12) than intermediate-low-proficiency readers' (M = .20, SD = .25), and were almost comparable to the explicit condition (M = .06, SD = .11). These low inference error rates suggest that inference ideas were activated to such an extent that high-proficiency readers responded to inference questions at the similar speed across the two explicitness conditions. Taken together, these results from high-familiarity texts provide the evidence that high-proficiency readers made better use of text content familiarity than did intermediate-low-proficiency readers, in their processing of inference generation during expository reading.

4.1.6.3 Difficulty with low-familiarity texts

The discussions made in the above two sections addressed the results obtained for high-familiarity texts. Nevertheless, to better understand the difficulty experienced by EFL readers with causal bridging inference generation, it is also important to investigate participants' on-line processes with low-familiarity texts. Analyses of the low-familiarity texts results reveled opposite patterns in causal inference generation in both proficiency groups. Specifically, there was no significant difference in participants' target reading times between the implicit and explicit conditions. Further, correct response times for inference questions were significantly longer in the implicit condition than in the explicit condition, indicating that inference ideas were less available when they were not stated in texts (i.e., the implicit condition) than when they were stated (i.e., the explicit condition). These results thus confirmed the difficulty associated with unfamiliar text reading. Researchers have proposed that unfamiliar text content impedes on-line reading processes, as it forces readers to focus on individual text elements by drawing cognitive resources from higher-level processing (e.g., Best et al., 2005; Coté et al., 1998). Even proficient L1 readers have been shown to experience difficulty with unfamiliar expository texts (e.g., O'Reilly & McNamara, 2007; McNamara et al., 1996).

In particular, the most noteworthy finding regarding low-familiarity texts is that target reading times were not significantly different between the implicit and explicit conditions. The null effect of the explicitness factor on target reading times indicates that participants reading low-familiarity texts failed to notice causal gaps in target sentences in the implicit condition. This observation contrasts with the results observed for high-familiarity texts, which showed that both of the two proficiency groups noticed causal gaps in implicit-target sentences. These results for the low-familiarity texts suggest that processing what was explicitly present in low-familiarity texts demanded many cognitive resources. Past think-aloud studies have demonstrated that verbal protocols elicited from readers of unfamiliar expository texts are typically dominated by comments concentrating on the current sentence (e.g., Coté et al., 1998; McNamara, 2004). Due to such increased cognitive demands with the processing of the explicit text, it was difficult for participants to engage in higher-level conceptual processes to understand causal relations provided in target sentences. Collectively, the findings from low-familiarity texts suggest that the unfamiliarity of the text content poses a great challenge to EFL readers' expository comprehension processes, irrespective of their L2 reading proficiency.

4.1.7 Conclusion of Experiment 4

The goal of Experiment 4 was to reveal the conditions under which EFL readers make causal bridging inferences during reading expository text. There were three main findings regarding on-line generation of causal bridging inferences.

First, high-proficiency readers made causal bridging inferences on the condition that the content familiarity of the text was high. Second, intermediate-low-proficiency readers failed to make causal bridging inferences during expository reading, irrespective of the familiarity of the text content. Third, neither intermediate-low- or high-proficiency readers made causal bridging inferences during reading low-familiarity texts; participants seemed to exclusively focus on explicit text information, thus failing to even detect causal gaps in target sentences in low-familiarity texts. Taken together, these observations suggest that participants made causal bridging inferences during expository reading when L2 reading proficiency and the familiarity of text content were both high.

These findings converge to indicate that bridging inferences, which are critical for establishing in-text links, are relatively limited in EFL expository reading. In particular, when the content familiarity of text is low, even high-proficiency readers have difficulty interpreting explicit text ideas, let alone inferring implicit ideas from the text. Thus, it is likely that with unfamiliar expository text, comprehending causal relations among pieces of information, one of the key focuses of this dissertation, would be quite challenging for EFL readers.

It is important to note that expository texts used in Experiment 4 were quite short in length; they consisted of only four sentences in the explicit condition, and three sentences in the implicit condition. Such short text length made it impossible to fully explore whether and to how EFL readers made use of different numbers of causal relations to prior text to understand currently read information. In order to more directly examine the relation between EFL readers' expository on-line processes and the causal structure of expository text, I designed Experiment 5.

4.2 Experiment 5

4.2.1 Purposes and research questions

Experiment 5 drew on the causal network model of comprehension to examine the extent to which EFL readers' on-line processes and off-line memory reflect the causal structure of the expository text. To compute the number of causal connections each text statement in the text has to others, I performed the causal network analysis on an experimental expository text (Trabasso & van den Broek, 1985; Trabaso et al., 1989; Varnhagen, 1991). I then investigated effects of the number of causal connections on reading times for, and recall rates of those statements. Furthermore, I considered explicitness of causal relations in the text and readers' L2 reading proficiency as variables.

The major reason for considering the explicitness of causal relations is that this factor is likely to affect the relation between reading times and the number of ECCs (i.e., causal connections to earlier text). To be more specific, when causal relations are made explicit in a text (i.e., the explicit condition), reading times for any given statement should be facilitated (i.e., becoming faster) as a function of the number of ECCs. The main reason for this phenomenon is that causal relations are relatively readily available to readers in this case, as those relations are directly signaled by the text. This situation (i.e., reading times becoming faster as a function of the number of ECCs) is analogous to findings of previous narrative studies (Magliano et al., 1999; Radvansky et al., 2014).

On the other hand, when causal relations remain implicit in text (i.e., the implicit condition), reading times for statements are predicted to take longer as a function of the number of ECCs. This prediction is because integrating current statements with prior text in the implicit condition is assumed to involve readers inferentially searching for relations between them, as those relations are not explicitly provided (Mulder, 2008;

Myers et al., 1987). It is therefore likely that, in the implicit condition, reading times inflate for statements with higher numbers of ECCs, as increased numbers of ECCs should involve readers' more active searching for causal relations to prior text.

Given these predictions regarding on-line expository processes, I placed the first focus of Experiment 5 on the interaction between the number of ECCs and explicitness of causal relations (hereafter, the ECCs × Explicitness interaction) in reading times for statements. Furthermore, given that the Experiment 4 showed that the success of on-line inference generation depended on readers' L2 reading proficiency, the increased reading times associated with higher numbers of ECCs in the implicit condition might be limited in readers with high L2 reading proficiency.

As to recall, participants' recall performance may be impacted by the total number of causal connections (TCCs), not ECCs, that statements have to all the information in a text (Trabasso & van den Broek, 1985; Ushiro et al., 2010). Literature has shown that ideas that are connected to a whole text are strongly encoded and better remembered later than ideas that are not. I explored whether this finding is replicable when explicitness of causal relations is manipulated in expository text. The following research questions (RQs) were addressed in Experiment 5.

RQ5-1: Do the number of ECCs and explicitness of the causal relations interactively affect EFL readers' reading times for the expository text?

RQ5-2: Do the number of TCCs and explicitness of the causal relations interactively affect EFL readers' memory for the expository text?

4.2.2 Method

4.2.2.1 Participants

Fifty-two Japanese undergraduate and graduate students at University of Tsukuba took part in this experiment. Their majors were various, including literature, education, humanities, psychology, social studies, agriculture, engineering, and international studies. None of them participated in the previous experiments. Twenty-nine of them were female, and the other 23 were male. Their ages ranged from 18 to 27 years (*M* = 20.29, *SD* = 1.50). All of them had studied English for more than six years.

Their self-reports indicated that participants had overall English proficiency of the CEFR A2 to B2 levels, which were generally equivalent to Experiment 4's participants' (Council of Europe, 2001; Tannenbaum & Wylie, 2008, 2013); the TOIEC listening and reading test (*M* = 638.99, *SD* = 111.09, range = 405 to 920) and the EIKEN test (Grade 4 to Grade 1: Grade 4, *n* = 2; Grade 3, *n* = 2; Grade Pre-2, *n* = 3; Grade 2, *n* = 12; Grade Pre-1, *n* = 6; Grade 1, *n* = 1). The self-reports were collected by a paper questionnaire. Forty-four participants reported at least one of the scores, and the other eight participants reported none of the scores.

As in Experiment 4, participants were grouped by means of a median split (*Mdn* = 13) into the intermediate-low- and high-proficiency groups, according to their performance on the same L2 reading proficiency test as in Experiments 2–4 (*M* = 12.90, *SD* = 4.56, *Min*/*Max* = 5/22). Twenty-three participants formed the intermediate-low-proficiency group (*M* = 8.65, *SD* = 2.10, *Min*/*Max* = 5/11; the TOEIC listening and reading test: *M* = 520.21, *SD* = 60.33, range = 405 to 650), while the other 29 formed the high-proficiency group (*M* = 16.28, *SD* = 2.79, *Min*/*Max* = 13/22; the TOEIC listening and reading test: *M* = 760.07, *SD* = 63.23, range = 645 to 920).

Participants were further grouped into those who read the explicit text and those who read the implicit text. Thus, there were, in total, four participant groups, crossing 2 (Proficiency) × 2 (Explicitness) conditions (see Table 4.7). A 2 (Proficiency: high, intermediate-low) × 2 (Explicitness: explicit, implicit) ANOVA on L2 reading proficiency test scores confirmed that the high-proficiency group scored better than the intermediate-low-proficiency group, $F(1, 48) = 119.07$, $p < .001$, $\eta_p^2 = .71$. Neither the main effect of the explicitness, $F(1, 48) = 1.74$, $p = .194$, $\eta_p^2 = .04$, nor the Proficiency × Explicitness interaction was significant, $F(1, 48) = 0.08$, $p < .779$, $\eta_p^2 = .01$.

Table 4.7

*L2 Reading Proficiency Scores of Participants in Experiment 5 (*N *= 52)*

Proficiency	*M*	95% CI	*SD*	*Min*	Max
			Explicit condition		
ILP (*n* = 12)	9.00	[7.55, 10.45]	2.56	4	11
HP (*n* = 14)	16.86	[15.77, 17.94]	2.07	14	20
			Implicit condition		
ILP (*n* = 11)	8.27	[7.39, 9.15]	1.49	5	10
HP (*n* = 15)	15.73	[14.06, 17.41]	3.31	13	22

Note. ILP = intermediate-low proficiency; HP = high proficiency; CI = confidence interval.

4.2.2.2 Material

L2 reading proficiency test. The reading test from the previous experiments was employed. It consisted of 26 items from the pre-first and second grades of the EIKEN test (Obunsha, 2005a, 2005b). The reliability of the test was acceptable, with a Cronbach's α of .81.

Text and explicitness manipulation. I employed the expository passage used in Experiments 1 and 2. In this experiment, the explicitness of causal relations was manipulated (see Appendix D).

The explicitness manipulation followed the procedures specified by Ozuru et al. (2010). First, a text with implicit causal relations was created by removing connectives (e.g., *however*, *because*, and *therefore*) and replacing nouns with pronouns. A text with explicit causal relations was then created by taking the original text and (a) adding connectives, (b) replacing pronouns with corresponding noun phrases, (c) adding nouns to enhance argument overlap between sentences, and (d) adding explanatory information to clarify relations between adjacent sentences.

Two raters conducted the explicitness manipulation; when necessary, discussion was held with an English native speaker to ensure that expressions and discourse contexts were natural. Table 4.8 presents key text features. As shown by the number of causal connectives, the manipulation successfully increased causal connectives in the explicit condition, as relative to the implicit condition. Flesch-Kincaid grade level indicates that the explicit text was more difficult than the implicit text; this is to be expected, because this index is computed based on word length and sentence length, both of which were increased by the manipulation.

Table 4.8

Features of the Explicit and Implicit Texts

	Explicit	Implicit
Number of causal connectives	7	0
Flesch–Kincaid grade level	8.24	6.72
Number of words	237	188

Note. Causal connectives include *therefore, because,* and *as a result* that specify a causal relation between two clauses; specifically, the explicit-condition text had three, two, and two of *therefore*, *because*, and *as a result*, respectively; Flesch–Kincaid grade level and the number of words were computed by Word for Mac 2016.

Causal network analysis. To determine the number of causal relations each statement in the experimental text had to others, three Japanese graduate students and I performed causal network analysis (Millis et al., 2006; Trabasso et al., 1989; Varnhagen, 1991). We first divided the experimental passage in the explicit and implicit version into 29 or 27 statements, respectively, with each statement corresponding to a subject-verb clause (Ushiro et al., 2010; Ushiro et al., 2015). For the causal network analysis, we then selected 22 statements that were commonly present across the explicit and implicit conditions and were relevant to the main theme of the passage. These statements specifically described a series of events occurring in the human body in space (statements in the last paragraph were not considered because they were not directly relevant to those events).

We identified the presence of causal relations among statements by means of a counter factual test. Specifically, we followed other the causal network model studies using the criterion of necessity (e.g., Ushiro et al., 2015; Varnhagen, 1991). This criterion assumes a causal relation between events A and B, in that if event A had not occurred, then event B would not have occurred (Trabasso & Sperrey, 1985); event A is necessary for event B to occur. Consider Statements 6 ("While in space, the body is not affected by gravity") and 7 ("Therefore, blood and water do not travel to the lower parts of the body, especially legs."), for example. If the human body is affected by gravity, then, it is possible that blood and water may descend to the lower body due to gravity. Therefore, we can assume a causal relation between these two statements.

Furthermore, we used two additional criteria to define causal relations. First, a causal relation was identified between those statements that involved mediating events for their causal link (Otero et al., 2004; Trabasso et al., 1989). Let us consider Statements 21 ("In addition, the decreased body water makes the heart pump less blood than normal.") and 24 ("As a result, the heart becomes smaller."). Between these two statements are mediating Statements of 22 and 23 ("Thus, the heart does not work" and "as hard as on Earth", respectively), which provide additional explanations for why statement 21 leads to 24. We identified a causal relation between Statements 22 and 24 because we can assume that when the heart pumps blood normally, the heart does not shrink. Second, we did not identify causal relations between those statements that deal with essentially the same event or co-occur in situations in the text. For example, Statements 3 ("how the zero gravity of space would affect humans.") and 6 ("While in space, the body is not affected by gravity") describe the same thing. Therefore, we determined that a causal relation does not apply between them.

Using these criteria, four raters (three Japanese graduate students and I) separately analyzed each of the 24 statements' causal relations with all the other statements in the text. Several discussions were then held to resolve discrepancies. We finished the causal network analysis when we reached agreement on causal relations for all the 22 statements.

Based on the causal network analysis results, I computed the number of causal relations each statement had with prior text (i.e., ECCs). By following Ushiro et al.'s (2015) procedure, statements were then grouped into low-ECC (with zero to two ECCs; $k = 8$), middle-ECC (with three to eight ECCs; $k = 7$), and high-ECC statements (with nine to 15 ECCs; $k = 7$). As for TCCs, I compute the total number of causal relations that any given statement had to all the other statements. Statements were grouped into low-TCC (with zero to four TCCs; $k = 6$), middle-TCC (11 to 12 TCCs; $k = 7$), and high-TCC statements (14 to 16 TCCs; $k = 9$). Appendix E shows the experimental text (the explicit condition) with the number of ECCs and TCCs.

4.2.2.3 Procedure

Participants were individually tested. At the beginning, the experimenter explained the general purpose and procedures of the study to participants and gained informed consent. Participants were assigned either the explicit or implicit text to start the reading session.

In the reading session, participants were orally and visually instructed on how to read the experimental text on a PC screen. They practiced reading with one expository passage. The practice and experimental passages were presented one statement at a time, left-aligned on a PC screen. After the signal "ready?", participants' pressing of the "yes" button on a response pad RB-730 (Cedrus, CA, USA) initiated the first statement. Participants were instructed to read for understanding so that they could complete post-reading tests (details about the tests were not given). When they thought the current statement was understood, participants pressed the "yes" button, which prompted the replacement of the current statement with the next one. Reading times for each statement were recorded with Super Lab 5.0 (Cedrus, CA, USA). After reading the experimental text, participants took a recall test. They were asked to write down all that they could remember about the experimental text in Japanese. The recall test was finished within 15 minutes. Finally, participants took the L2 reading proficiency test for 30 minutes. The average time to complete the experiment was 60 minutes.

4.2.2.4 Scoring and data analysis

Reading times. Reading times for statements were divided by the number of syllables in the corresponding statement to account for differences in statement length. Statement reading times were removed if they were less than 100 ms or three *SD*s away from the mean for a participant, resulting in the removal of 3.40% of the data.

Recall. In scoring recall protocols, four statements corresponding to the last paragraph were excluded from consideration to avoid any recency effect of recall (Ushiro et al., 2015). Of the remaining 24 statements, 22 that were common in the explicit and implicit texts were used for scoring. A point was given when participants correctly produced the gist of the statement (about two-thirds of the meaning). However, responses that were scientifically incorrect or inconsistent with the text were not credited. Two Japanese graduate students (including I) scored 30% of the data. The inter-rater agreement was 92%. After disagreements were resolved through discussion, I scored the remaining data.

4.2.3 Results

4.2.3.1 Reading times for text statements

Table 4.9 shows the means, standard deviations, and 95% CIs of reading times for statements as a function of the proficiency groups, the explicitness of causal relation, and the number of ECCs. In order to obtain an overall picture of the reading time results, I first computed correlations between statement reading times and the number of ECCs separately for the explicit and the implicit conditions. In the explicit condition, statement reading times and the number of ECCs were negatively correlated, $r = -44$, $p = .041$, meaning that there was a tendency for shorter reading times to be associated with increased numbers of ECCs. On the other hand, there was no significant correlation between statement reading times and the number of ECCs in the implicit condition, $r = -.10$, $p = 671$. Together, these correlation results imply that patterns of statement reading times were different between the explicitness conditions.

Table 4.9

Reading Times (in millisecond) as a Function of the Proficiency Groups, the Number of Earlier Causal Connections, and the Explicitness of Causal Relations

	Low ECC			Middle ECC			High ECC		
Explicitness	*M*	95% CI	*SD*	*M*	95% CI	*SD*	*M*	95% CI	*SD*
					ILP group (*n* = 23)				
Explicit	780.26	[497.97, 1065.54]	504.22	704.34	[484.50, 924.18]	388.55	605.15	[486.45, 723.85]	209.80
Implicit	631.96	[475.63, 788.29]	264.54	736.13	[584.55, 887.72]	256.51	576.23	[456.48, 695.99]	202.64
					HP group (*n* = 29)				
Explicit	583.44	[467.10, 699.77]	222.09	599.99	[438.66, 761.32]	307.99	519.14	[401.79, 636.50]	224.04
Implicit	609.76	[522.81, 696.70]	171.81	705.91	[568.98, 842.84]	270.58	592.74	[490.38, 695.11]	202.28

Note. ILP = intermediate-low proficiency; HP = high proficiency; ECC = earlier causal connections; CI = confidence interval.

Next, I subjected statement reading times to a 2 (Proficiency: high, intermediate-low) × 2 (Explicitness: explicit, implicit) × 3 (ECCs: high, middle, low) three-way mixed ANOVA to examine possible effects of and interactions among these three factors. The results found a significant Explicitness × ECCs interaction, which is illustrated in Figure 4.6, $F(2, 96) = 3.73$, $p = .028$, $\eta_p^2 = .07$. A main effect of ECCs was also significant, $F(2, 96) = 11.58$, $p < .001$, $\eta_p^2 = .19$. Conversely, a main effect of the explicitness factor was not significant, $F(1, 48) = 0.02$, $p = .891$, $\eta_p^2 = .01$. Likewise, the proficiency factor had no significant main effect nor did it interact with other factors, all *p*s > .100. These findings mean that, overall, L2 reading proficiency did not affect reading times. Therefore, data were collapsed across proficiency groups in subsequent analyses.

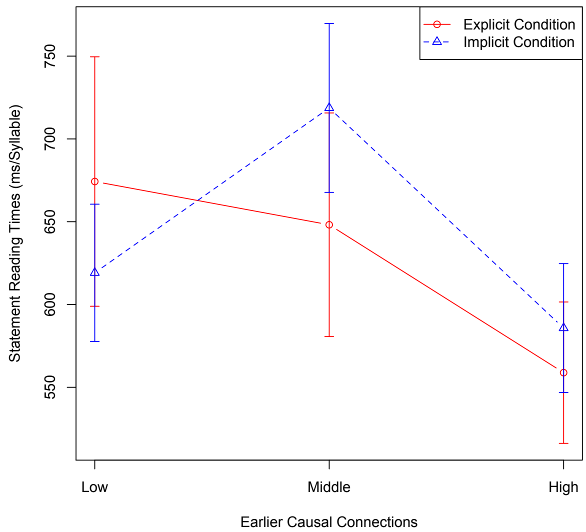

Figure 4.6. Statement reading times + error bars (standard errors) as a function of explicitness of causal relations and the number of ECCs.

To interpret the observed Explicitness × ECCs interaction, I performed follow-up tests. It was revealed that in the explicit condition, high-ECC statements were read significantly faster ($M = 558.84$, $SD = 217.67$) than middle-ECC ($M = 648.15$, $SD = 344.34$), $t(24) = 2.81$, $p = .031$, $d = 0.31$, and low-ECCs statements ($M = 674.28$, $SD =$

384.09), $t(24) = 2.50$, $p = .038$, $d = 0.38$. By contrast, reading times were not significantly different between low-ECC and middle-ECC statements, $t(24) = 0.93$, $p = .362$, $d = 0.08$.

In the implicit condition, longest reading times were found for middle-ECC statements ($M = 718.80$, $SD = 259.90$), compared to low-ECC ($M = 619.15$, $SD = 211.30$), $t(24) = 3.17$, $p = .010$, $d = 0.42$, and high-ECC statements ($M = 585.76$, $SD = 198.52$), $t(24) = 4.63$, $p < .001$, $d = 0.59$. Differences between low-ECC and high-ECC statements were not significant, $t(24) = 1.33$, $p = .195$, $d = 0.18$.

Finally, simple main effects of the explicitness factor were not found for any ECC statements, (all *p*s > .10), meaning that reading times were not generally affected by explicitness of causal relations in the text.

4.2.3.2 Recall rates

Table 4.10 shows the means, standard deviations, and 95% CIs of recall rates of statements, as a function of the proficiency groups, the explicitness of causal relations, and the number of TCCs. As in reading times, I first computed correlations between recall rates and the number of TCCs separately for the explicit and implicit conditions. The results showed that recall rates and the number of TCCs were positively correlated in the both explicit, $r = 51$, $p = .015$, and implicit condition, $r = .62$, $p = .002$. These positive correlations between recall rates and the number of TCCs propose that higher recall rates tend to be associated with higher numbers of TCCs, regardless of the explicitness of causal relations in the text.

Table 4.10

Recall Rates of Statements as a Function of the Number of Total Causal Connections and the Explicitness of Causal Relations

	Low TCC			Middle TCC			High TCC		
Explicitness	*M*	95% CI	*SD*	*M*	95% CI	*SD*	*M*	95% CI	*SD*
				ILP group (*n* = 23)					
Explicit	.22	[.12, .32]	.17	.36	[.30, .42]	.11	.58	[.47, .69]	.19
Implicit	.22	[.12, .33]	.18	.41	[.30, .51]	.18	.55	[.46, .63]	.14
				HP group (*n* = 29)					
Explicit	.20	[.09, .30]	.20	.41	[.31, .52]	.19	.59	[.51, .67]	.15
Implicit	.24	[.14, .33]	.20	.47	[.35, .58]	.23	.56	[.48, .64]	.16

Note. ILP = intermediate-low proficiency; HP = high proficiency; TCC = total causal connections; CI = confidence interval.

I subsequently subjected recall rates to a 2 (Proficiency: high, intermediate-low) × 2 (Explicitness of causal relations: explicit, implicit) × 3 (TCCs: high, middle, low) ANOVA to examine possible effects of these factors. The results revealed a significant main effect of TCCs, $F(2, 96) = 78.33$, $p < .001$, $\eta_p^2 = .62$. As shown in Figure 4.7, recall rates were significantly higher for high-TCC statements than for middle-TCC, $t(48) = 5.59$, $p < .001$, $d = 0.93$, and low-TCC statements, $t(48) = 13.77$, $p < .001$, $d = 2.06$. In addition, recall rates were significantly higher for middle-TCC statements than for low-TCC statements, $t(48) = 6.50$, $p < .001$, $d = 1.04$. It was therefore found that recall rates linearly increased from low- to high-TCC statements. No other effects or interactions reached statistical significance, all *p*s > .050.

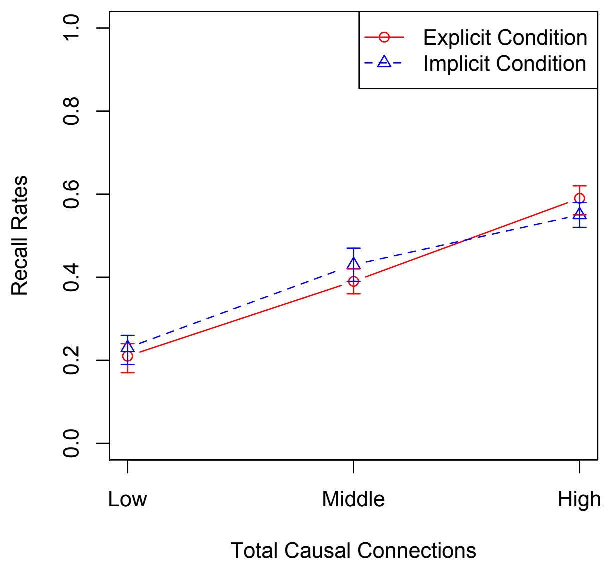

Figure 4.7. Statement recall rates ± error bars (standard errors) as a function of explicitness of causal relations and the number of TCCs.

4.2.4 Discussion

4.2.4.1 Effects of the ECCs × explicitness interaction on reading times (Research Question 5-1)

The results confirmed an Explicitness × ECCs interaction. On the one hand, participants in the explicit condition read high-ECC statements faster than middle-ECC and low-ECC statements. In line with this finding, statement reading times were negatively correlated with the number of ECCs. On the other hand, the longest reading times in the implicit condition were found for middle-ECC statements. A reliable correlation between statement reading times and the number of ECCs was not found. Note that this Explicitness × ECCs interaction held true for the both proficiency groups, as evidenced by the absence of any interaction associated with L2 reading proficiency. It should be noted that the null effect of L2 reading proficiency may seem inconsistent with Experimental 4's results, where only high-proficiency readers made causal bridging inferences. I addressed this issue in the subsequent experiment. Below, I provide separate discussions for the explicit and the implicit conditions to interpret the observed Explicitness × ECCs interaction.

Explicit condition. In the explicit condition, shorter reading times were associated with increased numbers of ECCs, which is in line with the causal network model (e.g., Magliano et al., 1999; Radvansky et al., 2014). The causal network model posits that processing of any given information is easier when that information has many ECCs than when that information has fewer ECCs (e.g., Langston et al., 1999). This facilitation is due to the fact that information related to what was understood before in a text can be more easily integrated with existing text mental representations than information that is less related (Radvansky et al., 2014). Therefore, the observed reading time facilitation for statements with higher numbers of ECCs indicates that participants processed incoming statements based on causal relations that those statements had to prior text. In the explicit condition, how each statement was causally related to others was made fully explicit with linguistic markers. Such explicitness might increase the availability of causal relations to participants for processing current statements, resulting in the found facilitation in reading times for high-ECC statements.

These results provide the evidence that the causal network model can be extended to on-line processes in EFL readers of expository text. In general, the causal structure in narratives consist of characters' goal-action relations that are embedded in one's daily lives (e.g., Graesser & Clark, 1985; Trabasso & van den Broek, 1985; Trabasso et al., 1989). In contrast, causal relations in the present expository text are scientific, physical, and less embedded in daily lives (see Britten & Gülgöz, 1991, for detailed discussion of features of the expository text). In this regard, the observed facilitation in processing as a function of the number of ECCs indicates that participants were, at least to some extent, sensitive to the scientific causal structure in the present text, on the condition that causal relations were made fully explicit.

Implicit condition. In contrast to the explicit condition, participants in the implicit condition took longer to read middle-ECC statements than other statements. The increased reading times indicate that participants engaged in additional cognitive processes to comprehend middle-ECC statements (Just & Carpenter, 1992). According to the causal network model, the prolonged reading times in the implicit condition are associated with readers' inferential searching for causal relations to prior text (Langston et al., 1999; Suh & Trabasso, 1993). Specifically, the model assumes that readers try to bridge current and prior text when there are causal connections between them (Magliano et al., 1999; Millis et al., 2006). Because causal relations were implicit, causally bridging pieces of text in the implicit condition should involve readers inferentially searching for implicit connections. Given this assumption, the increased reading times for middle-ECC statements in the implicit condition suggest the possibility that participants may have been searching for those statements' causal relations to prior text, resulting in the delay in reading times.

It is noteworthy that middle-ECC statements had the moderate number of causal antecedents (three to eight) in prior text—lower than high-ECC (nine to 15) and higher than low-ECC statements (zero to two). This moderateness of causal connections may have caused participants to engage in inferential searching. This is indirectly reminiscent of L1 studies showing that L1 readers took additional inference generation to understand statements with moderate strength of causal relations to preceding events (Golding et al., 1995; Myers et al., 1987). This account is furthered in Chapter 5 (General Discussion).

At the same time, one may argue that if an increase in reading times in the implicit condition is caused by readers' inferential searching for causal relations, then, longest reading times should appear for high-ECC statements with the highest number of causal relations to prior text; the higher the number of causal relations, the greater the necessity of searching processing (e.g., Millis et al., 2006). However, as described in the above results section, participants read high-ECC statements significantly faster than middle-ECC statements. These shortened reading times for high-ECC statements may be explained in terms of two perspectives. The first explanation is that reading times might be impacted by the serial position of statements in the text. Specifically, most high-ECC statements appeared in the later part of the text. It is well known that readers take more times to comprehend information presented in earlier text than information presented in later text (e.g., Millis, Simon, & Ten Broek, 1998). Given this observation, the late serial position might cause participants to read high-ECC statements in shorter times than other statements. This possibility can be addressed by computing correlations between statement reading times and their serial positions. However, the analysis found no reliable correlation between statement reading times and their serial positions in the explicit, $r = .13$, $p = .542$, or in the implicit condition, $r = .21$, $p = .351$. It is therefore safe to assume that the observed shorter reading times for high-ECC statements, compared to middle-ECC statements, cannot be simply explained by the serial position effect.

The second explanation is that the causal relations high-ECC statements shared with prior text were conceptually evident to participants—because high-ECC statements had such large numbers of ECCs, the extensiveness of causal relations overrode their explicitness (or their implicitness, more precisely), making their relation to prior statements evident to participants. As a result, participants integrated high-ECC statements into developing mental representations easily, without as much inferential searching as needed for middle-ECC statements. This possibility is supported by the structure of the experimental text. The expository text used in this study had a simple cause-effect structure, where any given statements almost always described the outcomes of the previous statements. This fact, along with the later serial positions of high-ECC statements, suggests that participants had processed a substantial number of causal antecedents to high-ECC statements at the point when they encountered those statements. Such a process might allow participants to expect that the previous statements were causes of the current statement. Consequently, high-ECC statements' causal relations were made conceptually evident, which in turn prevented reading times for those statements from being delayed in the implicit condition.

Of these two explanations, the latter one (high-ECC statements' causal relations to prior text were evident to participants) seems more plausible when considering the absence of correlations between statement reading times and serial positions. However, it must be noted that reading time data alone cannot provide strong support for this claim—reading times do not inform the actual content of reading processes employed by readers at a specific time point. Therefore, I further addressed this point in Experiment 6.

4.2.4.2 Effects of number of TCCs on recall rates (Research Question 5-2)

The results from the recall test revealed that participants increased recall rates linearly from low- to high-TCC statements. In line with this finding, recall rates were positively correlated with the number of TCCs. These findings together suggest that participants had better memory for text statements with higher numbers of TCCs than statements with fewer numbers of TCCs. This observation is in line with the causal network model (e.g., Langston et al., 1996; Trabasso et al., 1989), demonstrating that participants encoded the causal structure of the original text into their memory representations. This finding is a successful replication of Ushiro et al.'s (2015) results showing that EFL readers recalled information with many TCCs better than information with fewer TCCs in expository text.

Furthermore, the present experiment found that this linear recall increase by the number of TCCs was true for both the explicit and implicit conditions. This fact implies that even when causal relations between text statements were not made explicit in the text, participants' memory reflected the causal structure of that text. As stated in the discussion of the reading time results, cognitive processes involved in causally linking statements in the implicit condition are assumed to have been different from those processes in the explicit condition. In the explicit condition, causal relations between statements were made linguistically fully explicit. Therefore, what was required to detect causal relations to link statements was, to no small extent, comprehending the explicit text as instructed by linguistic markers. In contrast, causal relations between statements were not explicitly provided in the implicit condition. Accordingly, to causally link statements and capture causal structure in the text, participants would have needed to inferentially search for causal relations that potentially connect current and prior text. From this perspective, the fact that participants' memory reflected the text's causal structure in the implicit condition implies that participants detected implicit causal relations between statements.

In relation to this argument, one noteworthy finding from the reading time results is that participants in the implicit condition took the longest times to comprehend middle-ECCs statements (as discussed in the previous section). It is possible that during those additional processing times, participants searched for causal relations to prior text and encoded those statements into memory based on the inferred causal relations. As a result, their memory reflected the causal structure of the text. Though this argument may seem attractive, a strong claim cannot be made with this experiment's data alone because, as I mentioned in the previous section, reading times do not provide direct information about contents of processes. Thus, I complemented this possibility with a think-aloud method in Experiment 6.

4.2.5 Conclusion of Experiment 5

The goal of Experiment 5 was to explore how EFL readers' on-line processes and off-line memory reflect the causal structure of expository text. The following three major findings were derived.

First, participants' on-line processes reflected the causal structure of the text, when causal relations were made fully explicit (i.e., the explicit condition). Specifically, statements reading in the explicit condition were most facilitated for high-ECC statements, compared to other statements. This finding indicates that participants drew on causal relations to prior text to process current statements.

Second and contrastingly, participants showed the longest reading times for middle-ECC statements when causal relations were implicit in the text (i.e., the implicit condition). This suggests that participants engaged in additional cognitive processes to comprehend statements with moderate numbers of ECCs in the implicit condition.

Third, participants' off-line memory reflected the causal structure of the text, regardless of the explicitness of causal relations; they linearly increased recall rates as a function of the number of TCCs. Together, Experiment 5's findings suggest that EFL readers' on-line processes are more complexly related to the text's causal structure than their off-line memory. That is, the relation between on-line processes and the text's causal structure was modified by the explicitness of causal relations in the text.

I must note several points that remained unresolved. The first is in regard to contents of reading processes employed by EFL readers during expository comprehension. As noted in the discussion section, temporal data of reading times do not directly inform what processes participants engaged in during reading. Therefore, the findings from this experiment needs to be complemented by additional investigation of reading processes' contents. The second point is the relation between EFL readers' on-line processes and causal understanding of the expository text. Specifically, the process that may or may not contribute to EFL readers' understanding of causal relations cannot be explored by reading time data alone. In order to address these points, I designed Experiment 6 where participants' on-line processes were assessed by the think-aloud method.

4.3 Experiment 6

4.3.1 Purposes, hypothesis, and research questions

The goal of Experiment 6 was twofold: to reveal contents of on-line processes during EFL readers' understanding of causal relations in expository text, and to reveal the relation between EFL readers' on-line causal bridging and off-line causal understanding of expository text.

Regarding on-line processes, Experiment 5 provided two main findings about the self-paced reading method. First, participants' processing of high-ECC statements was facilitated, relative to other statements, when causal relations were made explicit in the text. This finding was indicated by shorter reading times for high-ECC statements than other statements. Second, participants were found to engage in additional processing to comprehend middle-ECC statements when causal relations were implicit in the text. This finding was indicated by middle-ECC statements having the longest reading times in the implicit condition. To complement these findings in this experiment, I used the think-aloud method, which can provide a window into the exact nature of comprehension processes engaged in by readers during reading (e.g., Ericsson & Simon, 1993). As in the previous experiment, I considered readers' L2 reading proficiency and the explicitness of causal relations in the text as variables.

Regarding the effect of L2 reading proficiency on verbal protocols. I hypothesized that L2 reading proficiency would have different effects on the frequency of mentioning currently read statements (termed *current statements*) and the frequency of causally bridging current and to prior sentences (i.e., causal bridging). To be more specific, I predicted that current statements would occur more frequently in less proficient readers than in proficient readers. On the other hand, causal bridging was expected to occur more frequently in proficient readers than in less proficient readers. This prediction is based on past studies' findings that less proficient readers often devote the majority of their cognitive resources to interpreting explicit meanings of current sentences, thus leaving few resources for interconnecting multiple sentences (e.g., Magliano et al., 2011; Millis et al., 2006; Shimizu, 2015). L1 research has also reported that skilled readers tend to actively causally bridge current and prior text to build coherence across different parts of text (e.g., Higgs et al., 2017; Magliano & Millis, 2003; Millis et al., 2006). Therefore, I hypothesized that the proportion of verbal protocols for current statements and causal bridging would cause an interaction with L2 reading proficiency.

In addition, this experiment investigated incremental patterns of causal bridging according to the number of ECCs. The causal network model holds that readers are more likely to engage in causal bridging for statements that have more ECCs than statements that have fewer ECCs (e.g., Langston et al., 1999; Magliano et al., 1999; Trabasso & Suh, 1993). I explored the applicability of this theoretical assumption to L2 expository comprehension.

Finally, Experiment 6 examined the relation between on-line processes and off-line causal understanding (assessed by the causal question). As Millis et al. (2016) demonstrated that L1 readers' causal bridging of current and distant prior information contributes to their causal understanding, I focused on whether EFL readers' causal bridging is correlated with their causal-question performance. The following hypothesis and research questions were addressed in this experiment:

Hypothesis: Proportions of verbal protocols for current statements and causal bridging interact with L2 reading proficiency.

RQ6-1: Does L2 reading proficiency affect incremental patterns of causal bridging as a function of the number of ECCs?

RQ6-2: Does the explicitness of causal relations in text affect incremental patterns of causal bridging as a function of the number of ECCs?

RQ6-3: Does causal bridging in EFL readers of different L2 reading proficiency relate to causal understanding (as assessed by a causal question)?

4.3.2 Method

4.3.2.1 Participants

Sixty-nine Japanese undergraduate and graduate students at University of Tsukuba took part in Experiment 6. Data from three participants were removed from analyses, as one of them had participated in the previous experiments, and the other two had technical trouble with recording verbal protocols. They were majoring in agriculture, education, humanities, international studies, literature, medicine, physics, psychology, and social studies. Thirty-one participants were female, and 38 were male. Their ages ranged from 18 to 29 years old (M = 21.01, SD = 3.16). All of them had studied English for more than six years.

According to their self-reports, their overall English proficiency was estimated to be basic to intermediate-advanced, or CEFR levels A2 to B2, which was almost equivalent to that of Experiments 4 and 5's participants (Council of Europe, 2001; Tannenbaum & Wylie, 2008, 2013); the TOIEC listening and reading test (M = 650.18, SD = 113.22, range = 405 to 940) and the EIKEN test (Grade 4 to Grade 1: Grade 4, n = 0; Grade 3, n = 6; Grade Pre-2, n = 9; Grade 2, n = 14; Grade Pre-1, n = 7; Grade 1, n = 1). The self-reports were collected by a paper questionnaire. Fifty-one participants reported at least one of the scores, and the other 18 participants reported none of the scores.

As in Experiments 4 and 5, intermediate-low- (n = 33, M = 7.97, SD = 2.47, Min/Max = 2/13; the TOEIC listening and reading test: M = 533.96, SD = 57.18, range = 405 to 600) and high-proficiency (n = 33, M = 18.79, SD = 3.27, Min/Max = 14/25; the TOEIC listening and reading test: M = 776.01, SD = 58.79; range = 600 to 940) groups were formed by means of a median split (Mdn = 13), based on their scores on the same L2 reading proficiency test as in Experiments 2–5(M = 13.38, SD = 6.16, Min/Max = 12/22).

Participants were also assigned to either an explicit or implicit condition, in which they read an explicit or implicit experimental passage, respectively. Table 4.11 shows experimental groups' L2 reading proficiency test scores.

Table 4.11

*L2 Reading Proficiency Scores of Participants in Experiment 6 (*N *= 66)*

Proficiency	*M*	95% CI	*SD*	*Min*	Max
		Explicit condition			
ILP (n = 16)	7.56	[6.24, 8.89]	2.71	2	12
HP (n = 16)	18.81	[17.05, 20.58]	3.60	14	25
		Implicit condition			
ILP (n = 17)	8.27	[7.29, 9.42]	1.49	5	13
HP (n = 17)	15.73	[17.32, 20.21]	3.31	14	25

Note. ILP = intermediate-low proficiency; HP = high proficiency; CI = confidence interval.

A 2 (Proficiency: high, intermediate-low) × 2 (Explicitness: explicit, implicit) ANOVA confirmed that the high-proficiency group scored better than the intermediate-low-proficiency group, $F(1, 62) = 225.66$, $p < .001$, $\eta_p^2 = .78$. Neither a main effect of Explicitness, $F(1, 62) = 0.27$, $p = .608$, $\eta_p^2 = .01$, nor the Proficiency × Explicitness interaction was significant, $F(1, 62) = 0.34$, $p = .563$, $\eta_p^2 = .01$.

4.3.2.2 Material

L2 reading proficiency test. The same reading test was used as in the previous experiments. The reliability of the test was acceptable, with a Cronbach's α of .81.

Text. The expository passage from Experiment 5 was used. Either an implicit or explicit version was presented to participants.

Target statements for thinking aloud. Following researchers' suggestion that having participants think every statement aloud may lead to fatigue (Millis et al., 2006) or statement-by-statement processing unreasonably (Caldwell & Leslie, 2010), this experiment was designed to elicit participants' verbal reports only from pre-determined target statements. Nine target statements were selected. Three of them were low-ECC statements (Statements 3, 5, 6), another three were middle-ECC statements (Statements 8, 11, 20), and the other three were high-ECC statements (Statements 15, 17, 24).

These target statements were selected because they were judged to be important for understanding the overall meaning of the experimental text in a norming study (Ozuru et al., 2010). In the norming study, 10 Japanese university students read an explicit experimental text for understanding. They were then asked to rate each statement in terms of importance to understanding the text's overall meaning, on a five-point scale (1 = *unimportant*, 2 = *rather unimportant*, 3 = *neither unimportant nor important*, 4 = *rather important*, 5 = *important*). As a result, three statements rated as particularly important to understanding were selected from each of the low-, middle-, and high-ECC statement groups. In the think-aloud task, these target statements were presented in red type.

Causal question. The same causal question was used as in Experiments 1 to 3 in Study 1.

4.3.2.3 Procedure

Participants were individually tested. First, the experimenter explained the research purpose and the overall procedure of the think-aloud session to participants. Informed consent was then obtained.

Training think-aloud session. This experiment had a think-aloud training session so that participants could familiarize themselves with the experimental procedure. The inclusion of the training session followed the suggestion from past research that valid data collection of verbal protocols requires participants to be sufficiently accustomed to verbalizing their thoughts as they read (Ericsson & Simon, 1993; Magliano et al., 1999). The present think-aloud training session consisted of three phases: (a) instruction, (b) demonstration, and (c) practice.

In the first instruction phase, the think-aloud instructions were presented orally as well as visually (on a PC screen) to participants. Participants were instructed to verbally report all thoughts that came to mind after reading each target statement. Following Ericsson and Simon's (1993) guidelines, participants were discouraged from overtly elaborating on what was being read. In line with past L2 think-aloud studies (e.g., Horiba, 2000), participants were also explicitly asked to report thought content. The specific instruction was as follows:

> There are several sentences colored in red in the text. After you read those red sentences, please verbally report all the thoughts that come to your mind when understanding those sentences. Be sure to not only read aloud or translate the current sentence, but verbalize "content" of your thoughts while understanding that sentence.

In the demonstration phase, the experimenter demonstrated how to think aloud with a separate passage. A passage was presented on a PC screen in a cumulative manner, such that previously read statements remained on the screen (Horiba, 2000; Shimizu, 2015), using PowerPoint for Mac (Microsoft, Tokyo). The experimenter silently read each statement for understanding and pressed the "Enter" key on a keyboard after understanding a statement. The key press initiated the next statement. For each target statement, the experimenter demonstrated thinking aloud in Japanese. After the demonstration, the experimenter asked if participants had any questions. Some participants asked questions about specific steps of thinking aloud (i.e., "Should I read aloud read sentences before saying thoughts?"). If necessary, the demonstration passage was again presented, and the experimenter repeated some part of the demonstration to address participants' questions.

In the practice phase, participants practiced thinking aloud with the passage used in the demonstration. Participants silently read each statement for understanding and thought aloud in Japanese after reading each target statement. If participants kept silent for 10 seconds or more after a target statement, or if a target statement was skipped accidentally, the experimenter reminded participants of the instruction to verbalize thoughts (e.g., "Do you have any thought for understanding the red sentence? Anything is OK"; Horiba, 2013). After participants finished the practice, the experimenter gave feedback. The feedback was often related to the specificity of verbal reports. The experimenter also asked whether participants had any questions, and answered them if they did.

Experimental think-aloud session. After the training session, participants read the experimental passage while thinking aloud, sharing their thoughts in Japanese, for each target statement. As in the training session, the experimental passage was cumulatively presented on a PC screen. As in the practice phase, participants were reminded of thinking aloud by the experimenter if they kept silent for 10 seconds or more after target statements, or if they skipped target statements. Verbal protocols in this experimental phase were recorded with an integrated circuit recorder.

After the think-aloud task was finished, a causal question was administered. Participants completed their responses to the causal question within 15 minutes. Finally, the L2 reading proficiency test was administered for 30 minutes. The average time to complete the experiment was 70 minutes.

4.3.2.4 Scoring and data analysis

Categorization of verbal protocols. Verbal protocols were transcribed from recorded audio by two raters. Protocols were subsequently parsed into statements, each equivalent to a subject-verb clause in Japanese.

Each statement was categorized as one of six types of verbal protocols: (a) paraphrasing, (b) local causal bridging (hereafter, *local bridging*), (c) distal causal bridging (distal bridging), (d) elaboration, (e) monitoring, and (f) others. The definitions and examples of each type of protocol are shown in Table 4.12.

Table 4.12

Categories of Verbal Protocols With Definitions and Examples

Category	Definition	Example
1. Current statements	The reader mentions the current statement.	*As a result, the heart becomes smaller than on the earth.*
2. Local bridging	The reader causally bridges the current statement and its causal antecedents within the two preceding statements.	*(The heart becomes smaller) because there is no need in space for the heart to work as hard as on the earth, I think.*
3. Distal bridging	The reader causally bridges the current statement and its causal antecedents that were three or more statements back.	*As the amount of the body water decreases, the heart sends less amounts of bloods. Because of this, the size of the heart becomes smaller.*
4. Elaboration	The reader elaborates on the current statement based on world knowledge.	*I imagine that when you do not walk, your legs will become thinner and weaker. And I think this is the same thing for the heart.*
5. Monitoring	The reader comments on the degree of their understanding.	*So, the heart becomes smaller. I understand.*
6. Others	The reader makes evaluative or emotional reactions or other mentionings.	*Really? it's surprising for me...the heart becomes smaller in space...*

Note. Examples come from verbal protocols for Statement 24 ("As a result, the heart becomes smaller.")

These six categories were pre-constructed based on this study's theoretical interests and research questions. The categories were selected following past expository comprehension studies using the think-aloud method (e.g., Magliano et al., 2011; Millis et al., 2006). Several points should be noted about the categorization. First, verbal protocols categorized as current statements could have been further split into sub-processes such as word form analysis and grammatical parsing—a form of sub-categorization that has been used often in past think-aloud studies with L2 readers (e.g., Horiba, 2000, 2013; Shimizu, 2015). However, because this study's research interest was on what participants pay attention to (i.e., current statements, causal antecedents, or world knowledge), not on how they process current information, I decided to treat participants' mentioned current statements as one protocol type.

Second, verbal protocols were categorized as causal bridging (distal and local causal bridging) only when participants used causal expressions (e.g., だから [because], その結果 [as a result]) to link current statements to prior text. On the other hand, verbal protocols were categorized as "others" when participants linked current statements to prior text by means of temporal succession (e.g., そして [and]). Furthermore, when links between statements were incorrect or incomplete, corresponding protocols were categorized as "others," even if participants tried to bridge current statements and prior text. Examples of such incorrect bridging protocols include ones that were based on participants' misunderstanding or ones that were produced in an incorrect cause-effect order (e.g., effects → causes).

Third, causal bridging was differentiated into two types, distal and local bridging, according to the distance between current statements and their causal antecedents. This categorization follows research by Millis et al. (2006), who showed that distal bridging, not local bridging, predicted L1 readers' performance on causal questions.

Two raters, including I, independently categorized 30% of the data. Inter-rater agreement was 89%, and disagreements were resolved through discussion. Using the refined criteria, I categorized the remaining data.

The total number of statements for each verbal protocol type was computed separately for low-, middle-, and high-ECC statements, by summing the number of low-, middle-, and high-ECC statements in each protocol. Proportions of each protocol type were then computed by dividing the number of each protocol type by the total number of verbal protocols within that ECC level (low, medium, or high). To illustrate, if a participant were to produce four protocols for current statements, two for local bridging, and one for elaboration for low-ECC target statements, the total number of verbal protocols for low-ECC statements would be seven. Therefore, the proportions of current statements, local bridging, and elaboration would be 57%, 29%, and 14%, respectively. Propositions were employed as the unit of analysis because this study's focus was relative amounts of attention devoted by participants toward different types of on-line processes.

Causal question. Performance on the causal question was computed using the same scoring system as in Experiments 1, 2, and 3. Two raters, including I, independently scored 30% of the data, yielding an inter-rater agreement of 93%. After disagreements were resolved through discussion, I scored the remaining recall protocols.

4.3.3 Results

The results are reported focusing on three aspects of findings that are relevant to the hypothesis and the RQs. First, I compare overall proportions of the six categories of verbal protocols between the two proficiency groups to address the hypothesis. Second, I report the results regarding incremental patterns of causal bridging from low- to high-

ECC statements to address RQ6-1 and RQ6-2. Third, I report the results regarding the relation between on-line causal bridging and causal understanding to address RQ6-3.

It must be noted that this experiment's statistical analyses focused on only factors that were most relevant to the hypothesis and RQs. This focus was adopted because there were so many factors in this experiment (i.e., L2 reading proficiency, types of verbal protocol, number of ECCs, and explicitness of causal relations), and considering all of them would have made interpreting the results unreasonably difficult.

4.3.3.1 Proportions of overall verbal protocols

First, I compared proportions of the six categories of verbal protocols between high- and intermediate-low-proficiency groups to address the hypothesis (Proportions of verbal protocols for current statements and causal bridging interact with L2 reading proficiency.). Analyses focused on differences associated with categories of verbal protocols and L2 reading proficiency. To perform the analyses, the data were collapsed across the three types of ECC statements and across the two explicitness conditions.

Table 4.13 shows the proportions of the six categories of verbal protocols in the two proficiency groups. In order to examine whether the two-proficiency groups differed in these proportions, the data were submitted to a 2 (Proficiency: high, intermediate-low) × 6 (Category: current statements, local bridging, distal bridging, elaboration, monitoring, others) ANOVA. The results showed a significant main effect of the category, $F(2.25, 144.17) = 153.99$, $p < .001$, $\eta_p^2 = .71$. However, this was qualified by a significant Proficiency × Category interaction, $F(2.25, 144.17) = 18.66$, $p < .001$, $\eta_p^2 = .23$, as shown in Figure 4.8.

Table 4.13

Proportions of Verbal Protocols as a Function of Verbal-Protocol Categories and the Proficiency Groups

	ILP group (*N* = 33)			HP group (*N* = 33)		
Category	*M*	95%CI	*SD*	*M*	95%CI	*SD*
Current statements	.62	[.55, .70]	.23	.40	[.36, .44]	.11
Local bridging	.10	[.07, .13]	.08	.25	[.22, .28]	.09
Distal bridging	.01	[.01, .02]	.02	.09	[.07, .12]	.08
Elaboration	.07	[.02, .11]	.13	.06	[.03, .09]	.10
Monitoring	.03	[.01, .05]	.06	.02	[.01, .04]	.04
Others	.17	[.13, .21]	.11	.17	[.14, .20]	.09

Note. ILP = intermediate-low proficiency; HP = high proficiency; CI = confidence interval.

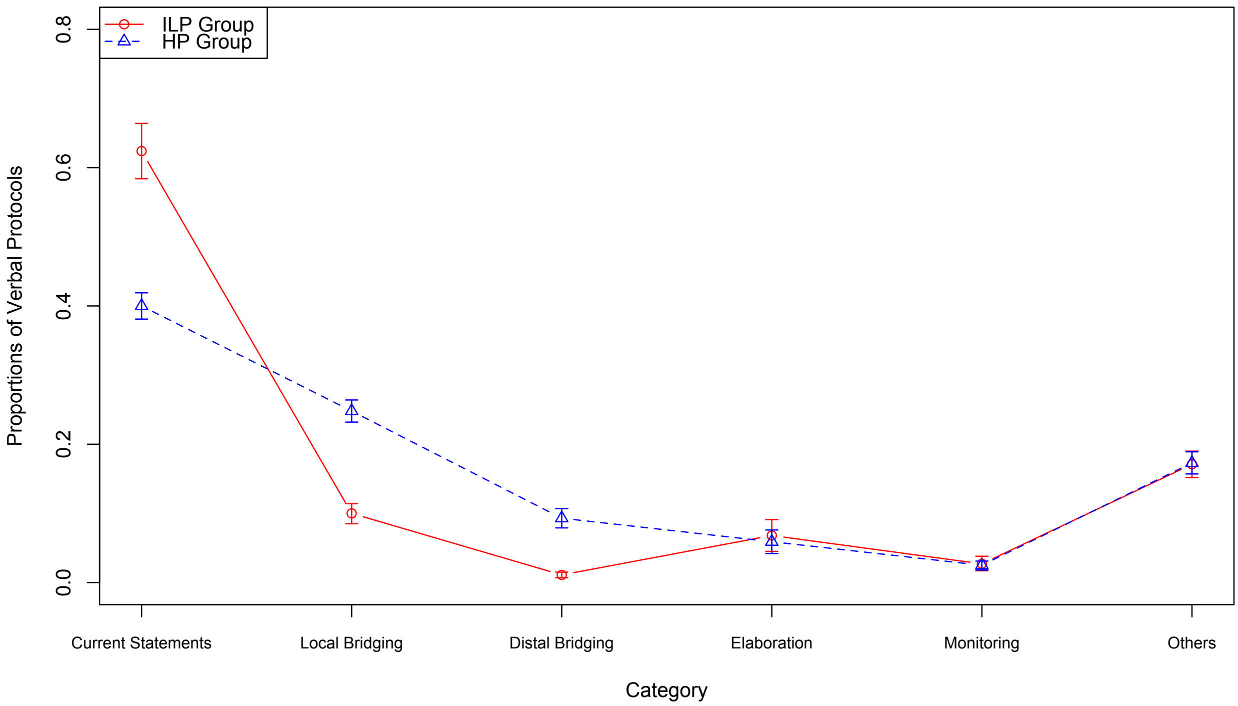

Figure 4.8. Proportions of the six categories of verbal protocols in the intermediate-low-proficiency (ILP) and the high-proficiency (HP) groups.

Follow-up tests on this Proficiency × Category interaction showed that intergroup differences existed only in proportions of current statements, local bridging, and distal bridging. The other categories of verbal protocols showed no significant group differences, all *p*s > .050. Specifically, proportions of verbal protocols for current statements were significantly higher in the intermediate-low- than in the high-proficiency group, $F(1, 64) = 25.57$, $p < .001$, $\eta_p^2 = .29$. On the other hand, proportions of verbal protocols for local bridging and global bridging were significantly higher in the high- than in the intermediate-low-proficiency group: local bridging, $F(1, 64) = 48.38$, $p < .001$, $\eta_p^2 = .43$; distal bridging: $F(1, 64) = 31.47$, $p < .001$, $\eta_p^2 = .33$. These results support the hypothesis, indicating proportions of current statements and those of causal bridging were oppositely affected by L2 reading proficiency.

4.3.3.2 Increase of causal bridging as a function of the number of ECCs

This section reports results regarding incremental patterns of causal bridging from low- to high-ECC statements, to address RQ6-1 (Does L2 reading proficiency affect incremental patterns of causal bridging as a function of the number of ECCs?) and RQ6-2 (Does the explicitness of causal relations in text affect incremental patterns of causal bridging as a function of the number of ECCs?). Prior to analyses, verbal protocols for local and distal bridging were summed to generate a comprehensive category, causal bridging. This step was taken because including the distance of causal bridging (local vs. distal) as a factor made the results too complex to interpret. Given that Experiment 5 showed that patterns of statement reading times differed between the explicit and implicit conditions, I conducted analyses considering the explicitness of causal relations.

Table 4.14 presents the two proficiency groups' proportions of verbal protocols for causal bridging for low-, middle-, and high-ECC statements in the two explicitness conditions.

These proportions were submitted to a 2 (Proficiency: high, intermediate-low) × 2 (Explicitness: explicit, implicit) × 3 (ECCs: low, middle, high) three-way ANOVA in order to examine whether participants' causal bridging increased from low- to high-ECC statements. The results found that the explicitness factor did not have a significant main effect or interact with ECCs, both *p*s > .050. These findings indicate that incremental patterns of causal bridging were not affected by the explicitness of causal relations. However, a two-way Proficiency × Explicitness × ECCs interaction was significant, $F(1.64, 101.36) = 3.55$, $p = .041$, $\eta_p^2 = .05$. To interpret this interaction, I performed follow-up tests separately on the data from the explicit and implicit conditions.

Table 4.14

Proportions of Causal Bridging as a Function of the Number of Earlier Causal Connections and the Proficiency Groups

	Low-ECCs Statements			Middle-ECCs statements			High-ECCs statements		
Proficiency	*M*	95%CI	*SD*	*M*	95%CI	*SD*	*M*	95%CI	*SD*
				Explicit condition					
ILP (*n* = 16)	NA	NA	NA	.09	[.03, .14]	.12	.19	[.10, .28]	.19
HP (*n* = 16)	.01	[−.01, .03]	.04	.37	[.29, .44]	.15	.55	[.45, .65]	.20
				Implicit condition					
ILP (*n* = 17)	NA	NA	NA	.03	[−.01, .06]	.06	.23	[.15, .31]	.18
HP (*n* = 17)	.08	[.03, .13]	.10	.31	[.24, .37]	.14	.48	[.39, .56]	.17

Note. ILP = intermediate-low proficiency; HP = high proficiency; CI = confidence intervals; Proportions of causal bridging were computed by summing up local and distal bridging.

Explicit condition. To examine incremental patterns of causal bridging as a function of the number of ECCs, a 2 (Proficiency: high, intermediate-low) × 3 (ECCs: low, middle, high) ANOVA was performed on the proportions of causal bridging. It was found that both the high- and intermediate-low groups increased causal bridging from low- to high-ECC statements, as indicated by a significant main effect of the ECCs, $F(1.72, 51.45) = 75.21$, $p < .001$, $\eta_p^2 = .72$ (see Figure 4.9).

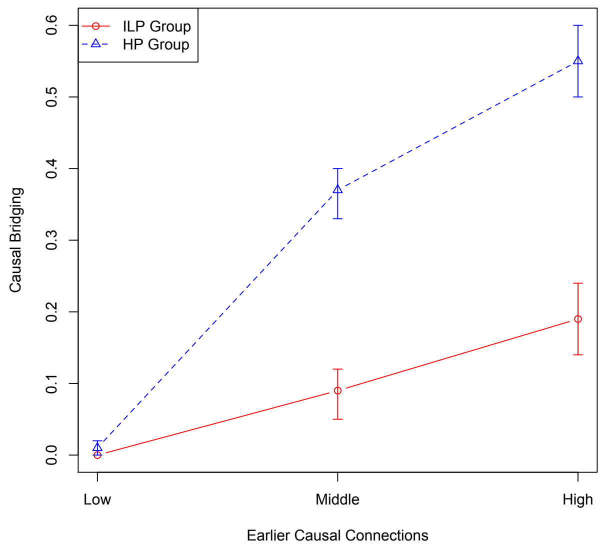

Figure 4.9. Proportions of causal bridging in the intermediate-low-proficiency (ILP) and high-proficiency (HP) groups in the explicit condition.

Specifically, the high-proficiency group implemented causal bridging most often for high-ECC statements, followed by middle-ECC statements and low-ECC statements: high- vs. middle-ECC statements, $t(15) = 5.37$, $p < .001$, $d = 1.03$; high- vs. low-ECC statements, $t(15) = 10.38$, $p < .001$, $d = 3.70$; and middle- vs. low-ECC statements, $t(15) = 9.20$, $p < .001$, $d = 3.25$. This linear increase in causal bridging was also found for the intermediate-low-proficiency group: high- vs. middle-ECC statements, $t(15) = 2.25$, $p = .040$, $d = 0.67$; high- vs. low-ECC statements, $t(15) = 4.04$, $p = .001$, $d = 1.43$; and middle- vs. low-ECC statements, $t(15) = 2.79$, $p = .014$, $d = 0.99$. Together, the results from the explicit condition show that participants as a whole increased causal bridging in a consistent way with the text's causal structure.

Implicit condition. As in the explicit condition, the 2 (Proficiency: high, intermediate-low) × 3 (ECCs: low, middle, high) ANOVA examined incremental patterns of causal bridging as a function of the number of ECCs. The results revealed a significant Proficiency × ECCs interaction, $F(1.5, 47.89) = 8.59$, $p = .002$, $\eta_p^2 = .21$, indicating that incremental patterns of causal bridging were different between the high- and intermediate-low groups (see Figure 4.10). Follow-up tests revealed that the high-proficiency group increased causal bridging linearly from low- to high-ECC statements, with the highest proportions for high-ECC statements, followed by middle-ECC statements and low-ECC statements: high- vs. middle-ECC statements, $t(16) = 4.37$, $p < .001$, $d = 1.08$; high- vs. low-ECC statements, $t(16) = 11.03$, $p < .001$, $d = 2.74$; middle- vs. low-ECC statements, $t(15) = 7.64$, $p < .001$, $d = 1.82$.

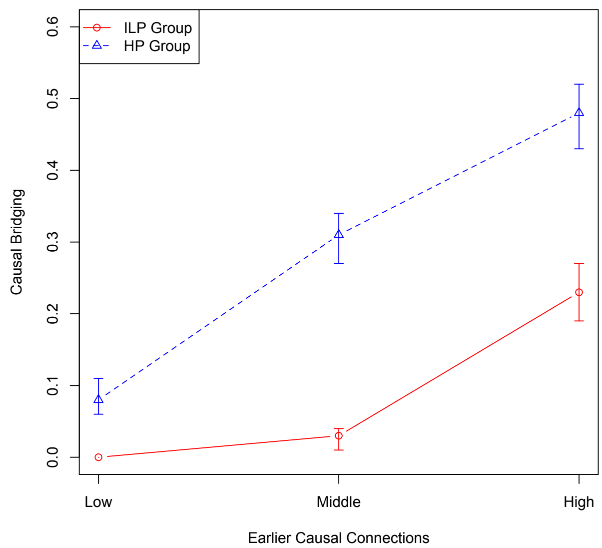

Figure 4.10. Proportions of causal bridging in the intermediate-low-proficiency (ILP) and high-proficiency (HP) groups in the implicit condition.

On the other hand, the intermediate-low-proficiency group failed to increase causal bridging from low- to middle-ECC statements, $t(16) = 1.85$, $p = .167$, $d = 0.63$, though their proportions of causal bridging were higher for high-ECC statements than middle-, $t(16) = 4.59$, $p < .001$, $d = 1.52$, and low-ECC statements, $t(16) = 5.37$, $p < .001$, $d = 1.84$.To be more specific, only 24% (4 out of 17) of the intermediate-low-proficiency group in the implicit condition showed causal bridging for middle-ECC statements. This rate was significantly lower than in the high-proficiency group, where all participants (17 out of 17) showed causal bridging for middle-ECC statements, $\chi^2(1) = 21.05$ $p < .001$, $\varphi = 0.79$.

This intergroup difference in causal bridging for middle-ECC statements is worth consideration—in Experiment 5, both the high- and intermediate-low-proficiency groups showed the longest reading times for middle-ECC statements, suggesting that participants as a whole engaged in additional processing. It is reasonable to assume that there were some processing problems for intermediate-low-proficiency readers when they read middle-ECC statements. Following this observation, the two Japanese graduate students (including I), inspected contents of intermediate-low-proficiency readers' verbal protocols for middle-ECC statements, in order to qualitatively examine how they processed middle-ECC statements.

As a result, two characteristics were found for the intermediate-low-proficiency group. First, 41% (7 out of 17) of the intermediate-low-proficiency group exclusively concentrated on current statements (i.e., produced almost only verbal protocols of current statements). This rate was significantly lower for the high-proficiency group, where no participants showed such a tendency, $\chi^2(1) = 8.81$, $p = .003$, $\varphi = 0.51$.

Second, the intermediate-low-proficiency group was found to have a tendency to incorrectly bridge middle-ECC statements and prior text. This tendency was observed in

41% (7 out of 17) of the intermediate-low-proficiency group and 6% (1 out of 17) of the high-proficiency group, $\chi^2(1) = 5.88$, $p = .015$, $\varphi = 0.42$. Consider the verbal protocol for Statement 20 ([it becomes more difficult for the human body] to work normally) below.

Excerpt 1: Participant H. S. (implicit condition)

Statement 20: to work normally normallyていうのは通常通り動くことが，難しいっていうのは最初はこのworkが働くって思ったんですけど，
で，の血とか水が体の中の血とか水が減ったときに，水が減ってるのがあんまり感じないので，
やっぱりその，水がこう減ったときに
普段より動かすのが難しいのかなって思いました。

The underlined part shows that this participant incorrectly explained why the body does not work normally in space; in fact, the human body sensitively responds to the lack of fluids. Another participant submitted the following protocol.

Excerpt 2: Participant I. S. (implicit condition)

Statement 20: ...で，体が水分をむしろ上に引き上げようとしてる，満たそうとするのに
水分量が減るから，バランスが，あの，おかしくて
人間の体が普段の働きをするのを難しくしているということです。

In this case, the participant's misunderstanding is shown in the underlined part, which states that the body tries to draw up water in space, leading to the weakening of body functioning; in space, the body actually tries to *decrease* body fluids.

4.3.3.3 Relation between on-line causal bridging and off-line causal understanding

This section reports findings pertaining to the relation between on-line causal bridging and ensuing causal understanding. Analyses focused on effects of L2 reading proficiency and the explicitness of causal relations. Table 4.15 presents performance on the causal question as a function of the proficiency groups and the explicitness conditions.

Table 4.15

Causal-Question Performance as a Function of the Explicitness of Causal Relations and the Proficiency Groups

Proficiency	*M*	95% CI	*SD*
		Explicit condition	
ILP (*n* = 16)	2.50	[6.24, 8.89]	1.26
HP (*n* = 16)	4.50	[17.05, 20.58]	1.41
		Implicit condition	
ILP (*n* = 17)	2.76	[7.29, 9.42]	0.75
HP (*n* = 17)	3.94	[17.32, 20.21]	1.39

Note. ILP = intermediate-low proficiency; HP = high proficiency; CI = confidence interval.

As in previous experiments, the high-proficiency group was found to perform significantly better on the causal question than the intermediate-low-proficiency group, $F(1, 62) = 27.43$, $p < .001$, $\eta_p^2 = .31$. No other effects reached significance (all *p*s > .050).

Next, to explore the relation between causal bridging and causal understanding, I computed correlations between proportions of local and distal bridging and causal-question performance in each of the four participant groups (which crossed 2 [Proficiency] × 2 [Explicitness] conditions). Table 4.16 shows the results.

Table 4.16

Correlations Between Local and Distal Bridging and Causal-Question Performance

	Explicit condition			
	Local bridging		Distal bridging	
Proficiency	*r*	*p*	*r*	*p*
ILP (*n* = 16)	−.32	.222	.11	.692
HP (*n* = 16)	.30	.253	.57*	.022
	Implicit condition			
	Local bridging		Distal bridging	
	r	*p*	*r*	*p*
ILP (*n* = 17)	.26	.315	.38	.133
HP (*n* = 17)	−.10	.577	.48*	.047

Note. **p* < .05; ILP = intermediate-low proficiency; HP = high proficiency.

As can be seen there, high-proficiency readers' distal bridging was significantly correlated with causal-question performance in the both explicit and implicit conditions. These results mean that the high-proficiency group was likely to answer the causal question better when they frequently implemented distal bridging than when they did less frequently. On the other hand, intermediate-low-proficiency readers' causal bridging (both local and distal) failed to be correlated with causal-question performance.

4.3.4 Discussion

4.3.4.1 Interaction between L2 reading proficiency and verbal protocols for current statements and causal bridging (Hypothesis)

The results confirmed the hypothesis, showing that proportions of current statements and causal bridging were affected by L2 reading proficiency in opposite directions. On the one hand, proportions of current statements were significantly higher in intermediate-low- than in high-proficiency readers. On the other hand, proportions of local and global bridging were significantly higher in high- than in intermediate-low-proficiency readers. These trends are generally consistent with past L2 studies that reported that less skilled L2 readers often exclusively concentrate on currently read information (e.g., Horiba, 1996; Zwaan & Brown, 1996).

Specifically, the present experiment found that 62% of verbal protocols by intermediate-low-proficiency readers were occupied by current statements, with causal bridging (the sum of local and distal bridging) constituting only 11% of their protocols (*SD* = 10%). This finding means that intermediate-low-proficiency readers devoted the majority of their attention (as indexed by their proportions of verbal protocols) toward what they were currently reading, leaving little attention for causal bridging. Indeed, there were some instances in intermediate-low-proficiency readers where they ended their verbal reports by only mentioning current statements. To illustrate, consider the following verbal protocol:

Excerpt 3: Participant A. S. (implicit condition)

Statement 15:	水，水飲まなくなっちゃう。
Statement 17:	体の水分レベルが普通よりも下がっちゃう。
Statement 20:	普通の動きをしなくなっちゃう。
Statement 24:	心臓は小さくなる。

This participant just repeated what was stated in current statements, without linking it to what had previously been read in the text. This excerpt shows that this participant's attention went no further than current statements. Another participant in the intermediate-low-proficiency group submitted the following protocol:

Excerpt 4: Participant S. S. (implicit condition)

Statement 15:	宇宙，宇宙旅行者はあまり水を飲まない。いや，飲む量が減る？ ああ、less, lessが形容詞とかにかかっていると訳しやすいんですけど, 水, って名詞にかかってたんで, 一瞬日本語訳が思いつきづらかったです。
Statement 17:	体内の水の水準が普段よりも低くなる で, これはどこから訳そうかっていうのを, 一瞬迷いましたけど，普通に主語から行けたんで。
Statement 20:	普段，普段通りに機能する。
Statement 24:	し，心臓が小さくなる。

In addition to repeating current statements, this participant tried to translate Sentences 15 and 17 into L1 (the underlined parts). Presumably, this participant's standards of coherence were set on interpreting the explicit meaning of current statements. Taken together, these protocols from two intermediate-low-proficiency participants illustrate the fact that they concentrated on explicit information from current statements and rarely made links across different parts of text.

On the other hand, the grand average of proportions of current statements in high-proficiency readers was 40% (*SD* = 11%). This finding indicates that high-proficiency readers managed to process currently read information with less attention than intermediate-low-proficiency readers. Instead, causal bridging constituted 34% of their verbal protocols (*SD* = 13%), showing that high-proficiency readers often used causal relations to prior text to understand current statements. For example, one participant in the high-proficiency group produced the following protocol for Statement 15 (“and space travelers drink less water”):

Excerpt 5: Participant M. T. (implicit condition)

Statement 15: はい，they drink less water. 水をほとんどspace travelersは飲まないとありますが，
その，それよりも前だと，その上の方に水が溜まってしまうので
ああ，heart and lungs send messages，体がメッセージを，何かしらメッセージを送ると。
で，どういうメッセージかというと，the amount of blood and water in the upper part of the body must be reduced. ということで，この上部の，こう，溜まった水や血液っていうのを減らそうとするメッセージというのが送られる。
なので，space travelers do not feel thirsty and they drink less water, という
space travelers, 宇宙飛行士，ん？宇宙旅行者っていうのは、喉の渇きが無いし、喉も乾かない。
なので，they drink less water. 水をほとんど飲まないっていうことに繋がっています。

The underlined part shows that this participant tried to causally explain the current statement (Statement 15) by drawing on distant information from Statements 11 (“like the chest and head are filled with blood and water”) and 12 (“the heart and lungs send messages”), in addition to recent information from Statement 14 (“space travelers do not feel thirsty”). This instance of distal bridging indicates that this participant sought to establish causal relations among pieces of information spread across different parts of the text. Such causal-reasoning processes are exactly what is postulated by models of discourse processing established in L1 reading research (e.g., Graesser et al., 1994; Magliano et al., 1999). Because high-proficiency readers processed current information relatively efficiently (as reflected by their lower proportions of current statements), they might have had more cognitive resources available for establishing causal coherence between different parts of text. As a result, they were able to implement distal as well as local causal bridging, in line with the discourse models.

4.3.4.2 Increase in causal bridging as a function of the number of ECCs (RQ6-1, RQ6-2)

Regarding RQ6-1 and RQ6-2, proportions of causal bridging were found to be a function of the interaction between L2 reading proficiency, the explicitness of causal relations, and number of ECCs. This section discusses the findings from this interaction by separately focusing on the explicit and implicit conditions.

Explicit condition. In the explicit condition, both high- and intermediate-low-proficiency readers increased causal bridging linearly as a function of the number of ECCs (i.e., high- > middle- > low-ECC statements). This linear increase indicates that regardless of L2 reading proficiency, participants' on-line processes reflected the causal structure of the text. This incremental pattern of causal bridging is also consistent with Experiment 5's reading time results that showed that high-ECC statements were most easily integrated with developing mental representations. It was accordingly corroborated that that when causal relations were made fully explicit in the text participants benefited from increased ECCs in their on-line expository comprehension processes.

At the same time, it must be noted that the results for intermediate-low-proficiency readers need careful interpretation. As described in the previous section, proportions of causal bridging were generally much lower in intermediate-low- than in high-proficiency readers. This fact proposes that the observed linear increase in intermediate-low-proficiency readers' causal bridging was only a relative one; intermediate-low-proficiency readers implemented causal bridging only relatively more often for statements with many ECCs than for statements with fewer ECCs. In other words, this increase does not absolutely mean that intermediate-low-proficiency readers actively engaged in causal bridging for statements with many ECCs.

Implicit condition. In contrast to the explicit condition, high- and intermediate-low-proficiency readers showed different incremental patters of causal bridging in the implicit condition. Specifically, high-proficiency readers increased causal bridging linearly from low- to high-ECC statements, consistent with the causal structure of the text. On the other hand, intermediate-low-proficiency readers did not increase causal bridging from low- to middle-ECC statements, in contrast to the text's causal structure. I thus provide separate discussions for high- and intermediate-low-proficiency readers below.

Regarding high-proficiency readers, their linear causal bridging increase means that they detected and drew on possible causal relations to prior text, even when causal relations were not made linguistically explicit. As discussed in the previous section, their high proportions of causal bridging, coupled with their low proportions of current statements, imply that high-proficiency readers' expository comprehension was guided by causal reasoning, by which they tried to explain current statements based on what had been previously understood in the text. It is assumed that during their causal reasoning, high-proficiency readers faced the necessity to search for implicit causal relations in order to explain current statements. This process resulted in their linear increase in causal bridging, in line with the causal structure of the text.

Particularly noteworthy is that high-proficiency readers implemented causal bridging significantly more for middle-ECC statements than for low-ECC statements. In Experiment 5, reading times in the implicit condition were longer for middle-ECC statements than for other statements. According to the causal network model, longer reading times in the implicit condition are associated with readers' inferential searching for causal relations to prior text (Langston et al., 1999; Suh & Trabasso, 1993). This experiment's results for high-proficiency readers match this view, indicating that they consciously tried to bridge middle-ECC statements and prior text by means of causal relations. As a result, their on-line processes reflected the causal structure of the text.

Unlike high-proficiency readers, intermediate-low-proficiency readers in the implicit condition failed to increase causal bridging from low- to middle-ECC statements. Qualitative inspection of verbal protocols provides two explanations for this finding. First, some of the intermediate-low-proficiency readers focused exclusively on currently read middle-ECC statements without implementing bridging processes. This trend suggests that those readers were, in the first place, not aware of implicit causal relations middle-ECC statements had to prior text. Second, intermediate-low-proficiency readers had trouble with the accuracy of bridging processes. Although some of them tried to bridge middle-ECC statements and previous parts of the text, they did so based on misunderstanding or incorrect reasoning. These patterns of verbal protocols combine to suggest that intermediate-low-proficiency readers had both quantitative (i.e., lack of bridging) and qualitative (i.e., inaccuracy of bridging) problems with causal bridging for middle-ECC statements in the implicit condition. As a result, they failed to increase causal bridging according to the number of ECCs.

Why did middle-ECC statement pose difficulty for intermediate-low-proficiency readers? A possible explanation is that those statements had moderate numbers of causal relations to prior text. Because there were not many causal relations between middle-ECC statements and previous parts of the text, correctly detecting how middle-ECC statements were related to prior text was demanding, especially for intermediate-low-proficiency readers, who tended to direct many of their attention to current statements. I provide an in-depth discussion of this argument in the general discussion section by additionally considering reading-time findings from Experiment 5.

Finally, it should be noted that high-ECC statements elicited the highest proportions of causal bridging in both the implicit and explicit conditions, which is reflected by the absence of an Explicitness × ECCs interaction. This finding means that upon encountering high-ECC statements, participants immediately detected those statements' implicit causal relations to prior text. They then used the detected causal relations to understand high-ECC statements, leading to the observed increase in causal bridging. This finding is in accordance with the reading time results in Experiment 5, which showed that processing high-ECC statements was easier than middle-ECC statements in the implicit condition. It seems that high-ECC statements' causal relations to prior text were conceptually evident to participants, which made it easy to causally integrate those statements with developing mental representations.

4.3.4.3 Relation between on-line causal bridging and off-line causal understanding (RQ6-3)

Regarding RQ6-3, causal understanding (as operationalized by performance on the causal question) was correlated only with the high-proficiency group's distal bridging. By contrast, neither of intermediate-low-proficiency readers' local nor distal bridging was correlated with causal understanding. The observed correlation between causal understanding and distal bridging partly parallels findings by Millis et al. (2006), who showed that distal bridging contributed to causal-question performance in L1 readers. At the same time, the present results stand in contrast to research by Horiba (2013), who failed to find any correlation between proportions of verbal protocols and recall rates in Japanese EFL readers. This discrepancy with Horiba's results may be explained by the fact that this study narrowly focused on causal relations in the text. Specifically, the analysis of this study targeted correlations between causal bridging and causal question performance, both of which are relevant to causal relations. On the other hand, Horiba analyzed correlations between all categories of verbal protocols and overall recall rates of experimental texts. Presumably, such highly specific focus of this study on text's causal relations exposed a link between on-line processes and off-line understanding more directly than in Horiba's study.

The correlation between high-proficiency readers' distal bridging and causal-question performance may be explained in terms of the large number of causal relations involved in distal bridging. As opposed to local bridging, which builds on causal relations between two consecutive statements, distal bridging involves establishing long-distance causal relations across different parts of text. This fact implies that engagement in distal bridging may allow readers to understand a large number of causal relations during reading. Then, because high-proficiency readers processed current information with a relatively few cognitive resources (as indicated by their lower proportions of current statements than intermediate-low-proficiency readers), they were supposedly able to allocate their cognitive resources to retain a large number of causal relations established by distal bridging. This process led to the observed correlation between distal bridging and causal understanding.

On the other hand, intermediate-low-proficiency readers devoted the majority of their cognitive resources to interpreting currently read explicit information. This focus on current statements is likely to leave few cognitive resources available for retaining causal relations built by local or distal bridging. As a result, their causal bridging failed to be correlated with off-line causal understanding.

However, we must be cautious of the fact that distal bridging hardly occurred in intermediate-low-proficiency readers. Specifically, only seven participants showed distal bridging in the intermediate-low-proficiency group, and this category of protocols accounted for only 1.1% of the intermediate-low-proficiency group's overall verbal protocols (SD = 2.2%). Such a small data size might be insufficient to yield statistically significant results. Therefore, the absence of correlations between causal understanding and distal bridging in intermediate-low-proficiency readers should be treated as tentative.

4.3.5 Conclusion of Experiment 6

In Experiment 6, I aimed to achieve two goals. First, I sought to reveal contents of on-line reading processes during EFL readers' understanding of causal relations in expository text. Second, I sought to reveal the relation between EFL readers' on-line causal bridging and off-line causal understanding. The findings can be summarized as three points.

First, causal bridging occurred significantly more often in high-proficiency readers than in intermediate-low-proficiency readers, whereas the opposite (intermediate-low- > high-proficiency readers) was true for verbal reports focusing on current statements. This finding indicates that high-proficiency readers made causal links between current and previous information, whereas intermediate-low-proficiency readers mostly concentrated on currently read explicit information. Second, incremental patterns of causal bridging by the number of ECCs was affected by the interaction between the explicitness of causal relations and L2 reading proficiency. In the explicit condition, both high- and intermediate-low-proficiency readers' causal bridging linearly increased from low- to high-ECC statements, consistent with the causal structure of the text. Conversely, in the implicit condition, intermediate-low-proficiency readers failed to increase causal bridging from low- to middle-ECC statements. Qualitative analysis of verbal protocols suggests that intermediate-low-proficiency readers in the implicit condition had the tendency to (a) focus on currently read statements and (b) incorrectly bridge current and prior information. Consequently, they failed to appropriately derive implicit causal relations between middle-ECCs statements and prior text. Third and finally, performance on the causal question was correlated only with high-proficiency readers' distal bridging.

Taken together, these findings indicate that all on-line processes and off-line causal understanding, and the relations between them, depend on L2 reading proficiency. High-proficiency readers routinely causally bridged different parts of text during expository reading, even when causal relations were implicit. They then built relatively robust causal understanding that was significantly correlated with their distal bridging. On the other hand, intermediate-low-proficiency readers interrelated different parts of text to a much lesser extent than high-proficiency readers. They showed weaker causal understanding, with which neither local or distal bridging was correlated.

4.4 Summary of Study 2

Study 2 explored on-line processes involved in EFL readers' expository text comprehension. First, Experiment 4 sought to reveal the conditions under which EFL readers make causal bridging inferences during expository reading. I considered L2 reading proficiency and the content familiarity of text and addressed RQ4, shown below.

RQ4: Under what conditions do EFL readers make causal bridging inferences during reading expository texts, when considering the content familiarity of the text and readers' L2 reading proficiency?

The results showed that EFL readers made causal bridging inferences when both L2 reading proficiency and the content familiarity of text were high. Readers with intermediate-low L2 reading proficiency, by contrast, failed to make appropriate inferences even from texts they were highly familiar with. For low-familiarity texts, furthermore, both high- and intermediate-low-proficiency readers focused on explicit text information, failing to detect causal gaps in target sentences.

Experiment 5 examined whether and how EFL readers' on-line processes and off-line memory reflect the causal structure of expository text. I considered the explicitness of causal relations in the text and addressed RQs 5-1 and 5-2, shown below.

RQ5-1: Do the number of ECCs and explicitness of the causal relations interactively affect EFL readers' reading times for the expository text?

RQ5-2: Do the number of TCCs and explicitness of the causal relations interactively affect EFL readers' memory for the expository text?

Regarding RQ5-1 (the relation between on-line processes and the causal structure), statement reading times were affected by the interaction between the number of ECCs and explicitness of causal relations. In the explicit condition, reading times were faster for statements with higher numbers of ECCs (high-ECC statements) than other statements, which is in line with the causal structure of the text. In the implicit condition, reading times were longest for middle-ECC statements, indicating that participants engaged in additional cognitive processes to comprehend these statements.

Regarding RQ5-2 (the relation between off-line memory and causal structure), participants showed a linear increase in recall rates as a function of the number of TCCs (high-TCC > middle-TCC > low-TCC statements), in both the explicit and implicit conditions. These results indicate that participants' off-line memory robustly reflected the causal structure of the expository text.

Finally, Experiment 6 explored contents of on-line processes involved in EFL readers' causal understanding of expository text. Experiment 6 also aimed to reveal the relation between on-line processes and off-line causal understanding. The following hypothesis and RQs were addressed.

Hypothesis: Proportions of verbal protocols for current statements and causal bridging interact with L2 reading proficiency.

RQ6-1: Does L2 reading proficiency affect incremental patterns of causal bridging as a function of the number of ECCs?

RQ6-2: Does the explicitness of causal relations in text affect incremental patterns of causal bridging as a function of the number of ECCs?

RQ6-3: Does causal bridging in EFL readers of different L2 reading proficiency relate to causal understanding (as assessed by a causal question)?

For the hypothesis, L2 reading proficiency affected proportions of current statements and causal bridging in predicted directions. Specifically, current statements occurred more frequently in intermediate-low- than in high-proficiency readers. By contrast, causal bridging occurred more frequently in high- than in intermediate-low-proficiency readers.

Regarding RQ6-1 and RQ6-2, incremental patterns of causal bridging were affected by the interaction between L2 reading proficiency and the explicitness of causal relations. In the explicit condition, both high- and intermediate-low-proficiency readers' causal bridging increased linearly from low- to high-ECC statements, which is consistent with the causal structure of the text. In the implicit condition, by contrast, only high-proficiency readers increased causal bridging according to the number of ECCs; intermediate-low-proficiency readers failed to increase causal bridging from low- to middle-ECC statements. Coupled with reading time results from Experiment 5, these verbal protocol findings indicate that EFL readers' on-line processes reflect the causal structure of expository text, either when causal relations are fully made explicit or when readers have high L2 reading proficiency.

Finally, regarding RQ6-3, performance on the causal question was correlated only with distal bridging in high-proficiency readers. The findings of the three experiments in Study 2 can be summarized in the following three points.

(1) EFL readers make causal bridging inferences during expository reading, when L2 reading proficiency and the content familiarity of text are both sufficiently high. (Experiment 4)

(2) EFL readers implement on-line reading processes reflecting the causal structure of expository text, either when L2 reading proficiency is high or when causal relations are fully made explicit (Experiments 5, 6)

(3) Only distal bridging in high-proficiency readers is related to causal understanding (Experiment 6)

Most past L2 studies on understanding of causal relations have used narrative text (Horiba, 1996; Ushiro et al., 2010). It remained therefore unclear how EFL readers process and understand scientific causal relations in expository text. Study 2's findings fill in this gap by demonstrating that EFL readers' on-line expository comprehension processes are complexly constrained by interactions between text and reader factors.

Chapter 5

General Discussion

This chapter combines findings from the six experiments in this dissertation to provide comprehensive discussion of EFL readers' causal understanding and learning from text. I specifically focus on the following three aspects of the findings, relevant to the three General RQs. First, I discuss the findings regarding the cognitive nature of EFL readers' causal understanding of expository text (General RQ1). Second, I discuss the findings regarding contributions of causal understanding to learning from text (General RQ2). Third and finally, I discuss the findings regarding how EFL readers' off-line memory and on-line processes reflect the causal structure the expository text (General RQ3).

5.1 Off-Line Memory and On-Line Processes Involved in Causal Understanding of Expository Text (General RQ1)

One main goal of this study was to reveal the cognitive nature of EFL readers' causal understanding of expository text. To this end, I explored off-line memory and on-line processes involved in EFL readers' causal understanding. Regarding off-line memory, the results from Experiment 1 found that the quality of text memory (i.e., how coherently text information is interconnected in memory) is more important for causal understanding than the quantity of memory (i.e., how much of information is recalled from text). Regarding on-line processes, Experiment 6 found that causal-question performance was significantly correlated only with distal bridging in high-proficiency readers. Taken together, these findings regarding causal understanding propose that EFL readers' causal understanding is associated with the specific processes. First, readers need to causally bridge pieces of information from different parts of text (i.e., distal bridging) during reading. Then, readers need to store the bridged pieces of information in an interconnected way in their memory representations.

In addition, this study found three conditions where causal understanding becomes difficult for EFL readers. First, causally bridging pieces of text becomes highly demanding, regardless of readers' proficiency, when the content familiarity of text is low (Experiment 4); it was shown that participants were forced to concentrate on current explicit information while reading low-familiarity texts. Second, retaining bridged pieces of text information becomes difficult when readers' L2 reading proficiency is intermediate-low or below (Experiment 6); the intermediate-low group devoted the majority of their attention to the processing of currently read information, which implies that they could not use enough attention to retain multiple pieces of information in an interconnected fashion. Third, just linking current and preceding information is insufficient for the attainment of causal understanding (Experiment 6); there was no correlation between causal understanding and local bridging (Experiment 6). To coherently understand text's causal relations, it is primarily necessary for readers to make links across different parts of the text so that they can establish global causal coherence of the text.

These findings jointly suggest that EFL readers' causal understanding can be achieved only when specific conditions are satisfied. That is, it is first necessary that the content familiarity of text is high; with low-familiarity text, relational processing is cognitively demanding even for skilled EFL readers. Then, a certain level of L2 reading proficiency (about the CEFR B2 level) is necessary so that readers can use their attention to both process and retain causal relations. Furthermore, readers need to make causal bridges across different parts of the text (i.e., distal bridging), not just between current and preceding text (i.e., local bridging).

It seems that when either of the above conditions are not met—such as when readers' proficiency or content familiarity of text is low—retaining pieces of information interrelatedly becomes highly difficult, preventing EFL readers from attaining causal understanding of text. This conclusion implies that EFL readers' causal understanding is much more limited than L1 readers'. It has been shown that L1 readers routinely try to causally bridge current and prior text during reading (Magliano et al., 1999; Trabasso & Magliano, 1996). Such causal bridging then contributes to their causal understanding of different genres of texts, such as narratives (Magliano & Millis, 2003), scientific expositions, and historical texts (Magliano et al., 2011; Millis et al., 2006). On the other hand, this study's findings demonstrate that EFL readers engages in causal bridging during expository reading under limited conditions; it is only when they sufficiently retain bridged pieces of information that causal understanding is built in EFL readers. Based on these cognitive characteristics of EFL readers' causal understanding, the next section discusses the relation between causal understanding and learning from text in EFL reading.

5.2 Contributions of Causal Understanding to Learning From Text (General RQ2)

This study aimed to reveal how causal understanding contributes to learning from text (as operationalized by performance on the problem-solving test) in EFL readers. The results found that causal-question performance contributed to problem-solving performance only in readers with high L2 reading proficiency (Experiment 2); low-proficiency readers' causal-question performance failed to predict their problem solving. It was thus evidenced that contributions of causal understanding to text learning depend on L2 reading proficiency (i.e., the Proficiency × Causal Understanding interaction in learning from text). This finding is important because there was almost no empirically grounded account of situation models in L2 readers. Specifically, although situation-model construction requires readers to go beyond text itself, past L2 studies exclusively used reproductive measures (e.g., recall, summary) that can be completed with the explicit text. This limitation was overcome in the present study by using the problem-solving test that required readers' application of knowledge to a new situation. Hence, the present study provides a new, empirically grounded account of EFL readers' situation models, showing that how those situation models relate to after-reading comprehension is modified by L2 reading proficiency.

This study also found that this Proficiency × Causal Understanding interaction in text learning is attributable to low-proficiency readers' trouble with both (a) understanding causal relations in text and (b) building accurate situation models of those causal relations. To be more specific, the experiments in Study 1 found four pieces of evidence.

First, low-proficiency readers constantly showed much lower performance on causal questions than did high-proficiency readers (Experiments 2, 3). Their poorer causal-question performance implies that they struggled to understand causal relations between pieces of text information. Second, low-proficiency readers exhibited significantly higher proportions of incorrect inferences in their answers to causal questions than high proficiency readers (Experiments 2, 3), which indicates that low-proficiency readers failed to build accurate situation models of causal relations in the text (Barry & Lazarte, 1998). Third, low-proficiency readers were found to have the tendency to answer the delayed problem-solving test with incorrect causal sequences and irrelevant information (Experiment 2), meaning that they could not achieve long-term learning of the text's causal relations. Fourth and finally, regardless of L2 reading proficiency, causal-question performance predicted performance on the problem-solving test when it was administered just after text reading (Experiment 3). This final evidence supports the view that the absence of a link between causal understanding and text learning in low-proficiency readers cannot be explained by the applied nature of the problem-solving test.

These four pieces of evidence converge to propose that low-proficiency readers had problems with processes at both textbase and situation-model levels. As a result, they could not acquire the text's causal relations as new knowledge. Given the fact that the situation-model construction usually builds on the appropriate understanding of the explicit text (Kintsch, 1998; McNamara & Kintsch, 1996), low-proficiency readers' difficulty with text learning is assumed to be rooted in the understanding of causal relations between units of information in the text. At this point, as discussed in the previous section, I found that causal understanding depends on the quality of text memory (i.e., how coherently bits of text information are interconnected in memory) more than the quantity of it (i.e., how much information is memorized from text). Therefore, the low-proficiency readers' problem is supposed to reside in processes with interconnecting relevant causal information to form networked mental representations. This view is indirectly supported by Experiment 6's finding that the intermediate-low-proficiency group rarely made causal bridges across different parts of the text (i.e., distal bridging).

Why did low-proficiency readers have trouble with interconnecting pieces of text? Two explanations are available from the present findings. The first relates to the limitation of cognitive resources. That is, low-proficiency readers may have to prioritize the processing of current information over connecting different parts of text in their allocation of cognitive recourses (Just & Carpenter, 1992), presumably due to their inefficient basic reading skills. This argument is supported by the fact that the intermediate-low-proficiency group devoted the majority of their attention to interpreting current statements

(Experiment 6). It is certain that the limitation of cognitive resources is at least partly responsible for low-proficiency readers' problem. However, this account alone seems insufficient to fully explain those readers' failure to interconnect text information, when considering that the experimental passages used in this study were all highly linguistically controlled; it is unlikely that such simplified passages caused so high cognitive demands.

The second explanation is that their standards of coherence were low. Specifically, low-proficiency readers were satisfied with understanding what they were currently reading, without causally relating it to prior text. A noteworthy finding is that the intermediate-low-proficiency group did not make causal bridging inferences even from high-familiarity texts (Experiment 4). Because these texts were highly linguistically controlled and short in length (three to four sentences), they were unlikely to impose so high cognitive demands. In other words, even under a cognitively undemanding situation, those readers did not make inferences necessary for understanding causal relations. This fact signifies that they placed their standards of coherence on the understanding of current explicit information, and thus, having little drive to interconnect pieces of information.

Note that that these two accounts (i.e., the limited-cognitive-resources view and the low-standards-of-coherence view) are not mutually exclusive. However, given the fact that low-proficiency readers concentrated on current information even with simplified texts, it might be reasonable to assume that their low standards of coherence were a root cause for their difficulty with causal understanding. This observation leads to the conclusion that their low standards of coherence, coupled with the limitation of cognitive resources, impaired low-proficiency readers' processes to interconnect relevant information, leading to their failure to understand the text's causal reactions.

In addition to considering processes associated with causal understanding, it is necessary to discuss situation-model-level processes; the observed Proficiency ×

Causal Understanding interaction in text learning indicates that low-proficiency readers could not retain the text's causal relations in situation models, even when they answered those causal relations (to the causal question) immediately after reading. At the situation-model level, less skilled readers were found to have both quantitative and qualitative problems with their inferential processing. First, regarding the quantitative problem, less skilled groups did not make inferences or engaged in inferential processing much less frequently than skilled groups across experiments (causal bridging inference generation in Experiment 4, inferential searching for prior causal relations in Experiment 6). This observation indicates that less skilled readers' on-line processes were mostly restricted to the textbase level.

Besides, in relation to the qualitative problem, less skilled readers had difficulty in their "accuracy" of inference generation. Specifically, low-proficiency readers constantly showed higher proportions of incorrect inferences in answering the causal question than high-proficiency readers (Experiments 2, 3). In addition, intermediate-low-proficiency readers inferentially bridged middle-ECC statements and prior text in an incorrect manner (Experiment 6). Apparently, less-skilled readers could not appropriately or correctly implement inferential processing, even when they tried to do so. Therefore, these qualitative and quantitative problems with inferential processing collectively lead to low-proficiency readers' failure to construct situation models of the text's causal relations.

This section discussed the Proficiency × Causal Understanding interaction in learning from text by focusing on the difficulties experienced by low-proficiency readers. Figure 5.1 summarizes their problems with textbase- and situation-model processes.

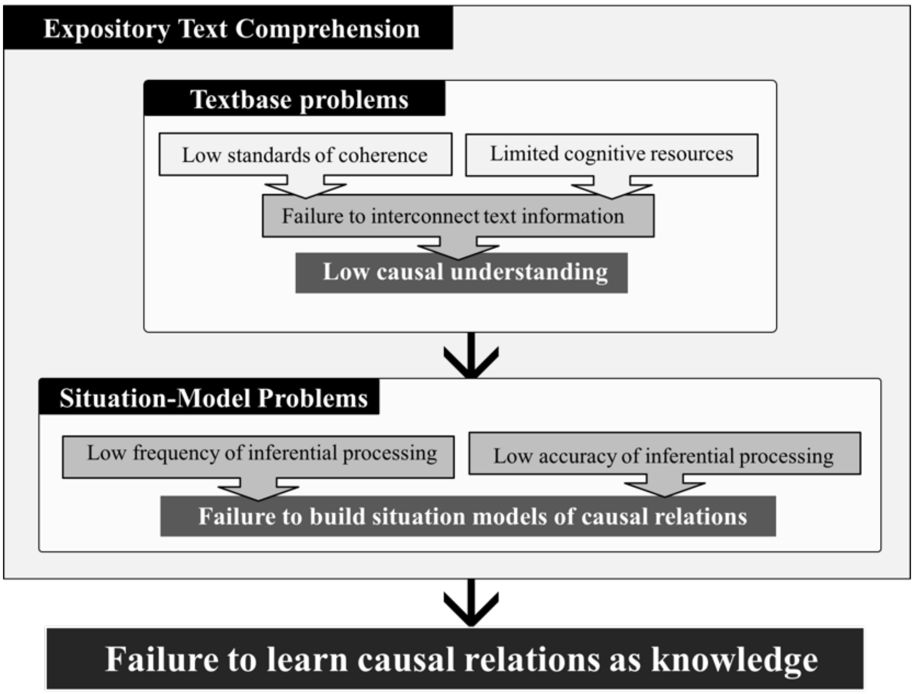

Figure 5.1. Low-proficiency readers' problems with learning from text.

At the textbase level, low-proficiency readers' processes were impaired by low standards of coherence coupled with the limited cognitive resources, leading to their failure to interconnect text information. These problems resulted in their low causal understanding. At the situation-model level, low-proficiency readers experienced both frequency and accuracy of their inferential processing, which resulted in their failure to build appropriate situation models of causal relations. These textbase- and situation-model-level difficulties jointly caused low-proficiency readers' failure to acquire causal relations as knowledge.

5.3 How On-Line Processes and Off-Line Memory Reflect the Causal Structures of Expository text (General RQ3)

The present study drew on the causal network model to explore whether and to what extent EFL readers' on-line processes and off-line memory reflect the causal structure of expository text (Experiments 5, 6). I tried to advance existing findings by looking at effects of a possible interaction between explicitness of causal relations and L2 reading proficiency on EFL readers' expository comprehension.

For off-line memory, the causal network model predicts that information with many TCCs is better recalled than information with fewer TCCs because information related to a larger part of text should be associated with increased retrieval access (Radvansky et al., 2014; Varnhagen, 1991). This prediction was supported by Experiment 5's result showing that participants increased their recall rates linearly from low- to high-ECC statements. Hence, participants' off-line memory reflected the causal structure of the original expository text, which is a successful replication of Ushiro et al.'s (2015).

In addition, the present study advanced Ushiro et al's (2015) findings by demonstrating that EFL readers' recall increase according to the number of TCCs held even when causal relations between text statements were implicit (e.g., in the implicit condition). This observation suggests that participants inferred implicit causal connections between text statements and encoded them into mental representations. Accordingly, this study proposes that EFL readers' memory is highly sensitive to the scientific causal structure.

As opposed to off-line memory, on-line processes involved in EFL readers' expository text comprehension were a function of the interaction between explicitness of causal relations and L2 reading proficiency. First, for the explicit condition (i.e., causal relations between pieces of information were made explicit), the causal network model predicts that an increased number of ECCs should facilitate the integration of current statements into evolving mental representations. This prediction was supported by the findings from the three-pronged approach. Reading times were faster for statements with highest numbers of ECCs (i.e., high-ECC statements) than for other statements (Experiment 5), meaning that participants tried to understand current statements based on their causal relations to prior text. This finding fits well with verbal protocol data in Experiment 6 revealing that participants' causal bridging of current and prior text increased linearly from low- to high-ECC statements in the explicit condition.

Therefore, when causal relations were explicit in the text, participants' on-line processes were consistent with the causal structure of the text. This conclusion matches past L1 studies that reported that L1 readers routinely draw on causal relations to prior text to comprehend incoming information (e.g., Magliano et al., 1999; Trabasso & Magliano, 1996). It may be that because the presence of causal relations was explicitly signaled by linguistic items, detecting how current and prior statements were related was relatively easy for participants. As a result, participants as a whole benefitted from increased ECCs in their on-line processes.

In contrast to the explicit condition, the causal network model predicts for the implicit condition that readers should engage in inferential searching for causal relations to prior text, which is associated with delayed reading times, when current statements have higher numbers of ECCs. This prediction was only partially supported in this study. Experiment 5 showed that, in the implicit condition, both high- and intermediate-low-proficiency readers took longer times to read middle-ECC statements than other statements. Experiment 6 then found that only high-proficiency readers successfully increased causal bridging according to the number of ECCs. Conversely, intermediate-low-proficiency readers failed to appropriately implement causal bridging when they read implicit-middle-ECC statements. Taken together, the findings in the implicit condition show that middle-ECC statements elicited inferential searching associated with longer reading times, and that only high-proficiency readers were successful in doing so.

A possible explanation for why inferential searching was performed for middle-ECC statements relates to the moderate numbers of ECCs that those statements possessed.

Specifically, because middle-ECC statements had moderate numbers of ECCs, they may have both "necessitated" and "allowed" inferential search. To understand this explanation, it is helpful to consider the other statements. First, high-ECC statements had an extensive number of ECCs, which may make high-ECC statements' causal relations to prior text conceptually evident. Therefore, high-ECC statements did not "necessitate" inferential searching by readers. This argument is supported by the fact that high-ECC statements in the implicit condition took significantly shorter reading times (Experiment 5) but elicited increased amounts of causal bridging (Experiment 6). The combination of facilitated reading times and increased causal bridging implies that participants readily identified high-ECC statements' causal relations without needing to take time to overtly search for those relations.

On the other hand, low-ECC statements had only a few numbers of ECCs. Because there were a few (if not no) causal relations to prior text, low-ECC statements hardly "allowed" participants to engage in inferential searching. In line with this, low-ECC statements in the implicit condition were associated with significantly shorter reading times than middle-ECC statements (Experiment 5) and elicited little causal bridging (Experiment 6). This pattern of results (shorter reading times with little causal bridging) implies that participants paid little attention to causal relations to prior text when they read low-ECC statements. In sum, high-ECC statements did not necessitate, and low-ECC statements did not allow, inferential searching for causal relations. Only middle-ECC statements that had moderate numbers of ECCs both necessitated and allowed inferential searching. Then, as shown in Experiment 4, such inferential processing was difficult for readers whose L2 reading proficiency was low; readers with sufficient L2 reading proficiency successfully implemented the processing.

Note that this argument (moderate numbers of ECCs elicited inferential searching) is analogous to the existing L1 finding that sentence pairs connected by moderately strong causal relations elicited bridging inferences more than strongly related and unrelated pairs (e.g., Myers et al., 1987). Though the strength of causal relations is not simply interchangeable with the number of causal connections, it is noteworthy that some L1 narrative studies have found that, like strong causal relations, increased numbers of causal connections have a facilitation effect on sentence reading times (Golding et al., 1995; Magliano et al., 1999; Radvansky et al., 2014). That is to say, reading times were similarly facilitated when the current information had strong causal relations and when current information had many causal connections to prior text. In essence, the processing of any text information is facilitated equally when the causal relation between current and prior text is quantitatively (i.e., the number of causal connections) or qualitatively (i.e., the strength of causal relations) higher. A definitive conclusion cannot be drawn from the present findings alone. However, it is at least evident in the present case that participants in the implicit condition faced the necessity of inferential searching when reading statements with moderate numbers of ECCs, as was the case for participants presented with moderately related pairs in the prior study (Myers et al., 1987). Clearly, additional research is necessary to better understand how commonly and differently the strength of causal relations and the number of causal connections influence EFL readers' on-line processes.

The present on-line findings jointly suggest differences between L1 and L2 readers' expository comprehension processes. In the context of L1 reading, many discourse models posit that deep-level text comprehension is guided by readers' causal reasoning, in which readers explain current information based on prior text and their prior knowledge (Graesser et al., 1994; Millis et al., 2006). Indeed, there is considerable evidence, both with narrative and expository text, that L1 readers routinely make causal bridges between pieces of text during comprehension (Magliano & Millis, 2003; Magliano et al., 2011). Unlike L1 readers, in EFL reading, processing individual sentences draws many of the cognitive resources from relational processing (Morishima, 2013; Zwaan & Brown, 1996). In addition, as discussed in the previous section, less skilled L2 readers tend to settle for the interpretation of individual text ideas, with their standards of coherence set on explicit pieces of text. Hence, it seems that only when the reader and text factors allow the reader (a) to direct a sufficient amount of cognitive resources toward causal relations and (b) to set standards of coherence on the understanding of causal relations that EFL readers make causal bridges between information in accordance with the text's causal structure. Specifically, in this study, when causal relations between pieces of information were made linguistically fully explicit, or when readers' L2 reading proficiency was sufficiently high, participants implemented causal bridging in line with the causal structure of the text.

At the same time, this argument implies that when conditions are satisfied, EFL readers can causally process and understand text input, which supports the assumptions in theoretical models (e.g., Graesser et al., 1994; Singer et al., 1992). Indeed, the high-proficiency group's processing was generally guided by causal reasoning; they made causal inferences explaining texts' causal relations during reading (Experiment 4) and tried to explain current statements based on causal relations to prior text (Experiment 6). In addition, when causal relations were explicitly signaled by linguistic markers, participants generally processed current statements using causal relations (Experiments 5, 6). These findings converge to suggest that, when supported by reader and text factors, EFL readers have the potential to strive for causally coherent mental representations, just as L1 readers do.

The participants' (conditional) tendency of causal processes is in line with the notion that causal relations constitute a fundamental part of text comprehension (Graesser & Clark, 1985; León & Peñalba, 2002; Trabasso et al., 1989). It seems that factors that made participants' causal processes difficult were related to L2 reading (e.g., inefficient language skills, the limitation of cognitive resources); when such L2-related problems were resolved (e.g., when readers had high L2 reading proficiency, or when causal relations were made explicit), participants generally processed the expository text causally. On this account, at least in the present experimental environments, the limited availability of causal relations in EFL readers is considered to be a problem of L2 language proficiency, rather than one of general cognitive abilities.

Taking the on-line and off-line findings together, this study revealed that the causal network model is differently applicable to EFL readers' on-line processes and off-line memory. On the one hand, the causal network model was highly predictive of participants' off-line memory for the expository text, as indicated by the linear increase in recall according to the number of TCCs. On the other hand, the causal network model only partly predicted participants' on-line expository comprehension processes. Because readers are under less cognitive pressure in the off-line task than in the on-line task, detecting and representing causal relations into memory representations might not be so cognitively demanding, especially compared to immediately processing causal relations during reading. In support of this claim, researchers have reported that readers may make inferences that were not generated during reading (Hosoda, 2014; Noordman et al., 1992).

At the same time, it must be noted that the observed strong sensitivity to causal structure in participants' off-line memory may be attributable to the nature of the recall test. In the recall test, participants were asked to retrieve what had been understood from the text. During this retrieval phase, participants may have reflected on how each bit of text information was connected to other information. Thus, the recall test allowed participants a "second chance" to pay attention to causal connections between text statements, leading to the observed recall increase according to the number of TCCs. In order to clarify this possibility, additional research is needed that include participants thinking aloud during the off-line task. Such an attempt would provide a better understanding of the difference between on-line processes and off-line memory in terms of sensitivity to causal relations in text.

Finally, this study showed that the three-pronged approach is a useful way to examine EFL readers' on-line processes. Whereas self-paced reading showed delayed reading times for implicit-middle-ECC statements for both high- and low-proficiency readers, the results from verbal protocols showed that low-proficiency readers had specific processing trouble with causal bridging for those statements. These findings support the notion that complementing temporal data with verbal protocols leads to a clearer understanding of moment-by-moment processes. At the same time, it must be noted that if the present study had employed the think-aloud method alone, I could not have excluded the possibility that low-proficiency readers' incorrect processing for implicit-middle-ECC statements was artificially caused by the nature of the think-aloud task (in which participants were asked to verbalize their thoughts). In this regard, the temporal data from self-paced reading are the evidence that participants paid more attention to implicit-middle-ECC statements than other statements. Therefore, the three-pronged approach has allowed for a comprehensive picture of EFL readers' on-line processes, with information about the both time course and contents of comprehension processes.

Chapter 6

Conclusion

6.1 Summary of Findings

The goal of this dissertation was to explore how Japanese EFL readers understand causal relations and learn from expository text. To achieve this goal, I addressed the following three General RQs.

General RQ1	What on-line processes and off-line memory are involved in EFL readers' causal understanding of expository text?
General RQ2	How does EFL readers' causal understanding contribute to their learning outcomes from expository text?
General RQ3	How do EFL readers' on-line processes and off-line memory reflect the causal structure of expository text?

Regarding General RQ1, the results show that high familiarity of text content is primarily necessary for EFL readers to process expository text causally. In addition, EFL readers need to have L2 reading proficiency of intermediate-high or above (the upper end of the CEFR B1 level or above) so that they can allocate their attention to relations between individual text elements. After these two conditions are met (i.e., both the text's content familiarity and the reader's L2 reading proficiency are sufficiently high), causal understanding can be achieved by readers' causally bridging different parts of the text and retaining them in an interconnected manner.

Regarding General RQ2, contributions of causal understanding to learning from text depended on L2 reading proficiency (Experiment 2, 3, 4). Specifically, causal understanding contributed to text learning in readers having intermediate-high or higher proficiency, but not in low-proficiency readers.

Finally, regarding General RQ3, EFL readers' off-line memory reflected the causal structure of the expository text, regardless of the explicitness of causal relations or L2 reading proficiency (Experiment 5). By contrast, EFL readers' on-line processes reflected the text's causal structure either when causal relations were made fully explicit or when readers' L2 reading proficiency was high (Experiments 5, 6).

Summarizing the above answers to the General RQs, we can conclude that the below three relations addressed in this dissertation are all dependent on L2 reading proficiency or on the interaction of L2 reading proficiency and text factors (i.e., the content familiarity of text, explicitness of causal relations).

With expository text whose content familiarity is not low:

(1) On-line processes and causal understanding

→Related only by high-proficiency readers' causal bridging

(2) Causal understanding and learning from text

→Related only in readers with proficiency of intermediate-high or above

(3) On-line processes and the causal structure of expository text

→Related only in high-proficiency readers or with causal relations made fully explicit

In L1 reading, theoretical models and empirical findings jointly underscore the importance of causal relations in the situation-model construction (Millis et al., 2006; van den Broek et al., 2002). The present findings jointly provide the evidence that in EFL reading, on-line processes, causal understanding, and text learning are not so straightforwardly related. It seems that these relations manifest themselves in EFL expository comprehension, only when the reader and text factors allow readers' standards of coherence to be placed on causal relations across different parts of the text (e.g., when L2 reading proficiency was high and causal relations were made fully explicit).

To date, most L2 reading studies have used narrative text. There is accordingly no theoretically sound or empirically grounded account of how EFL readers understand causal relations in expository text or of how they learn from text. This study fills in this gap and highlights cognitive characteristics of EFL readers' causal understanding and subsequent learning.

6.2 Limitations of This Study

The present study accordingly provides new insights into expository text comprehension processes in EFL readers. However, I must note that this study has several serious limitations that merit attention in future research. I focus on three limitations in this section: (a) the nature of the experimental texts, (b) the effects of prior knowledge, (c) differences in participants' proficiency between Studies 1 and 2, and (d) the effects of reading goals.

6.2.1 The nature of the experimental texts

Above all, it must be emphasized that most of the experiments in the present study (except for Experiment 4) employed only one highly simplified expository text. I believe that the present results, at least to some part, were due to the simplified nature of the experimental text. For example, the absence of a difference in recall rates between good and poor explainers in Experiment 1 is attributable to the high linguistic simplicity of the text; the limited amount of information made it easy for overall participants to encode and remember information from the text. Given the lack of comparison to texts of varying difficulties, it is difficult to generalize the present findings to more authentic expository text. It is recommended that future research uses several versions of authentic passages to explore the possible interplay of linguistic features of text and L2 proficiency in learning from text in L2 reading.

Taking the syntactic complexity for example, some L2 studies reported that syntactically complex passages elicit readers' elaboration or inferences more than less complex passages (Barry & Lazarte, 1998; Crossley & McNamara, 2016). Based on this finding, it is likely that readers with high L2 reading proficiency make more inferences and thus learn better from complex text than from less complex text. On the other hand, syntactically complex passages will likely force low-proficiency readers to devote most of their attention to analyzing structures of sentences. This would lead to low-proficiency readers experiencing greater difficulty learning from complex than from less complex text. In sum, L2 reading proficiency and syntactic complexity may interactively effect L2 readers' inferential processes and ensuing text learning. This possibility is an interesting topic to address in future studies.

6.2.2 The effects of prior knowledge

Second, this study did not consider the effects of participants' prior knowledge on participants' causal understanding and learning. It is widely agreed that how well readers learn from text is largely dependent on the amount of knowledge in the text's domain (Kintsch, 1994; McNamara et al., 1996). Experiment 4 considered the content familiarity of texts as a variable; however, the content familiarity in the present case was defined only based on participants' subjective perception (i.e., how familiar participants felt the text content was). In other words, this study did not assess the extent to which participants actually possessed knowledge on the text's subject matter. For example, a cold should be a common, familiar disease even to those who do not have medical knowledge on a cold.

In the present study's case, prior knowledge might have interacted with L2 reading proficiency. Some researchers have shown that struggling readers often fail to use or integrate prior knowledge with text information to deepen comprehension (e.g., McNamara & O'Reilly, 2009). Hence, high-proficiency readers were more likely to effectively use prior knowledge to learn better from the text. By contrast, simply possessing prior knowledge might have been insufficient to support low-proficiency readers' learning, especially considering that the present low-proficiency group struggled with understanding explicitly stated causal relations in the text (Experiments 2, 3). In general, effective knowledge use builds on the efficient operation of basic reading processes (Kintsch, 1998). It is likely that a certain level of L2 linguistic skills or knowledge is a prerequisite for making good use of prior knowledge to facilitate learning. An important direction for future research is to examine how L2 readers of different linguistic proficiency may or may not benefit from prior knowledge in their expository comprehension and learning from text.

6.2.3 Differences in participants' proficiency between Studies 1 and 2

Third, there were differences in participants' L2 reading proficiency between Studies 1 and 2. Specifically, Experiments 2 and 3's participants were generally less proficient than Experiments 4–6's participants (see Table 3.5). This limitation makes it difficult to compare the results from the two Studies in parallel.

In particular, the low-proficiency group in Experiments 2 and 3 had the lowest proficiency, according to their L2 reading proficiency test scores and self-reports. Because Experiments 2 and 3 did not assess on-line reading processes, low-proficiency readers' difficulty with their on-line processes was not directly clarified. When considering their lower proficiency, it is possible that the low-proficiency group might have failed to increase causal bridging according to the number of ECCs, even when causal relations were made fully explicit. This reasoning is because the low-proficiency group were assumed to pay even more attention to processing current statements than the Study 2's intermediate-low-proficiency group. Thus, the low-proficiency group might have had less attention available for processing causal relations between text statements.

At the same time, it is unlikely that the low-proficiency group were able to implement processes that intermediate-low-proficiency readers could not (e.g., the generation of causal bridging inferences, the appropriate operation of causal bridging in the implicit condition). It is natural to assume that the difficulties experienced by the intermediate-low-proficiency group would have occurred to a larger extent in the low-proficiency group. Additional research including participants with a wider range of proficiency is needed to gain direct information about the difficulty experienced by low-proficiency readers. Such an attempt would reveal a broader picture of individual differences in EFL readers' learning from text.

6.2.4 The effects of reading goals

Fourth and finally, this study did not examine the effects of participants' reading goals. Reading is usually a goal-directed activity (McCrudden et al., 2007). For example, expository text is usually read for study, whereas narrative text is read for entertainment. It is widely recognized that such reading goals determine readers' standards of coherence (Linderholm & van den Broek, 2002; van den Broek et al., 2002); when one reads for study, their standards for measuring coherence should be placed on learning what the text tries to communicate. This fact raises the possibility that low-proficiency readers, whose standards of coherence were low, read the text under a different reading goal than did high-proficiency readers. Specifically, low-proficiency readers were likely to focus on trying to understand the explicit meanings of individual sentences, rather than understanding the causal relations between them, as suggested by their focus on explicit information.

Additional research is needed to obtain a systematic account of how differences in reading goals lead to different degrees of expository text comprehension and text learning in L2 readers. It would also be significant to explore whether readers, whose default reading goal is to understand the explicit text, may or may not be supported by the explicit provision of a reading goal for learning from text.

6.3 Implications for Educators

In the conclusion of this dissertation, I suggest implications for EFL reading instruction. In most educational settings, expository text plays a vital role in conveying new information to students. Still, there are only limited instructional efforts to develop EFL students' abilities to learn from text. To inform instruction that can help students with their causal understanding and subsequent text learning, I discuss four implications below.

6.3.1 Assessing and promoting causal understanding of expository text

The present study provides evidence that causal relations play an important role in reading comprehension. It is recommended that teachers provide sufficient instructional support so that students can causally understand text. To this end, the present study's findings suggest a three-step reading instruction, as described below.

First, teachers should assess students' causal understanding with causal question(s); to promote causal understanding, it is first necessary to identify the extent to which students build causal understanding. In Experiment 1, I found that the mere recall of individual facts from the text does not relate to the ability to causally explain the text (i.e., causal understanding). It is therefore recommended that teachers should assess students' causal understanding by having students actually explain the text rather than just confirming students' recall of important information.

Second, if students cannot explain the text well, teachers can ask students whether they remember causally important information from the text. Considering Experiment 1's results, it is likely that, even when students cannot produce a coherent explanation, they may hold each bit of relevant causal information in their memory. Students should thus be encouraged to pay attention to causally important information in a reflective manner.

Third, explicit training should be conducted where students integrate pieces of causal information into one coherent explanation of the text. Difficulty with causal understanding is assumed to reside in processes that interconnect relevant information. Hence, after directing students' attention to causally important information, teachers should instruct students to consciously interrelate pieces of that information so that they can build relational text understanding. For example, teachers can use a diagram depicting causal relations in the text, such as the one in Figure 6.1.

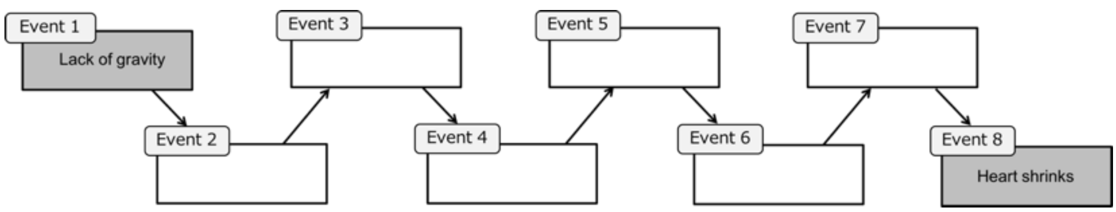

Figure 6.1. An example causal diagram.

Students are asked to complete the diagram by filling in the blanks with corresponding text events. In this task, the need to visually reconstruct the text's causal relations could draw students' focus to the information that is causally related to specific events. The deliberate implementation of this instruction will develop students' ability to better understand causal relations in expository text.

6.3.2 Interventions for supporting less-skilled students' expository comprehension

The present results clearly highlight a need for support for less skilled readers, whose English proficiency corresponds to the lower-end of the CEFR B1 level or below. Thus, this section specifically discusses intervention to support less-skilled readers' difficulty with expository text comprehension. The findings show that readers at this level may have difficulty with processes to relate explicit text information to each other; the low- or intermediate-low-proficiency groups struggled with (a) making causal links between current and prior text (Experiment 6) and (b) correctly answering causal relations to the causal question (Experiments 2, 3, and 6). Note that an intervention for lexical- or syntactic-level processes alone would be insufficient to assist less-skilled students in deepening expository comprehension, as suggested by the fact that the low-proficiency group still struggled with the highly linguistically simplified passage. Rather, the findings suggest that support for struggling EFL students should be aimed at facilitating processes involved in relational understanding.

One example is the provision of pre-reading questions that guide students' attentional allocation. Research has shown that questions specifically targeting important relations in text can make those relations explicit to students (McCrudden et al., 2009). For example, providing causal questions before reading would direct students' attention toward causal relations between information in text and in turn support the attainment of a causal understanding of that text. Actually, a number of L1 studies have proved that the provision of causal questions is a powerful way of supporting struggling readers' relational understanding (Carlson et al., 2014; McMaster et al., 2015).

6.3.3 Interventions for promoting skilled students' learning from text

For skilled students, the findings propose that an intervention aimed at enhancing distal bridging may deepen their causal understanding, which then would contribute to improving their text learning; in readers with English proficiency of the upper end of the CEFR B1 level or above, distal bridging was correlated with causal understanding (Experiment 6), which in turn contributed to long-term learning outcomes from text (Experiment 2). It should be noted that distal bridging accounted for only 9% of verbal protocols in the high-proficiency group (SD = 8%). This implies that even skilled students do not frequently engage in distal bridging on their own. Therefore, students should be explicitly encouraged to direct their attention toward causal information placed in distant parts of the text.

One example is the while-reading provision of the causal question that elicits students' explanation of the text (e.g., *Why does the body water level become lower in space?*). The causal question in this task must target events with causal antecedents in distant parts of the text so that students can build global causal coherence through explanation. By explicitly instructing students to explain causally important events, this method necessitates them to direct attention toward causal information in the distant text. L1 researchers often report that reading strategy trainings of this kind are indeed effective in promoting students' bridging processes (e.g., McNamara, 2004; Ozuru et al., 2010). Encouraging high-proficiency students' distal bridging in this way will deepen their causal understanding, which will subsequently allow them to learn better from text.

6.3.4 Applying the causal structure of expository text to reading instruction

This study revealed the effects of the causal structure of expository text on EFL readers' off-line memory and on-line processes. Two major pedagogical implications can be derived from this line of findings. First, text ideas having many causal relations to the other ideas in a text can be utilized as cues for post-reading activities. This implication comes from the fact that participants generally encoded high-TCC statements into memory robustly, suggesting that information with many TCCs may allow for effective retrieval of the text read. For example, after students read an expository text, a teacher can provide an event that has many causal antecedents and causal consequences in the text (such as "space travelers drink less water" in the present experimental text). Students are then asked to explain to each other what caused this event, and what consequences may result from it. Such a post-reading training might allow students to reflect and elaborate on the causal structure of the text read.

Second, interventions should be given for supporting and confirming students' understanding of information with moderate numbers of ECCs. This study found that, when causal relations remained implicit, middle-ECCs statements were difficult to process for readers with intermediate-low proficiency. Given the fact that causal relations in authentic expository text are not usually made explicit (Britten, 1994), teachers' support will likely play a significant role in students' successful causal processing. A theoretical discourse analysis, like the causal network model, can be used to identify where students will encounter processing difficulty in the text. Based on the information from the text analysis, students can be reminded of when or where they should pay attention so that they can effectively tackle the difficult parts of the text. Instruction informed by the theory of discourse analysis should be a valid way of supporting cognitive difficulty involved in EFL reading.

6.3.5 From "learning to read" to "reading to learn" instruction

In light of the English for academic purposes education, the present study's findings provide a theoretical basis for reading instruction for developing students who can learn necessary information from foreign sources and use it as new knowledge. Currently, English classes in Japan generally focus on grammar or translation, even at the university level. In other words, the instruction of "learning to read" has received much attention. I propose that the present study is a first step to advance this trend of English reading instruction in Japan by informing the direction of the "reading to learn." This study has revealed emerging difficulties as well as conditions necessary for EFL readers to learn from text.

Note that it is only when specific conditions are satisfied that EFL students can attain a causal understanding of and learn from expository text. This conclusion means that to develop students who can globally expand their expertise by learning from English text, intentional efforts are necessary so that environments can be established wherein students' learning is allowed. One important concept is standards of coherence. The findings propose that students should be encouraged to place their standards of coherence on understanding relations between pieces of the text, not just on individual text elements per se. When the conditions are satisfied—when the text's content familiarity and students' L2 reading proficiency is high, or when causal relations are explicitly signaled—the present participants causally interpreted, processed, and understood expository texts, with their standards of coherence set on causal relations between events in the text. It is therefore recommended that such conditions are intentionally arranged in the classroom. For example, causal relations should be made explicit by discourse markers when less-skilled students read expository texts, in order to reduce their cognitive burden at the textbase level. This will allow students to strive for a relational understanding of the text, without settling for individual words/sentences.

I also recommend that students be trained in the application of knowledge gained from the text, especially in college or university classes. For such applied instruction, the reading class can include the problem-solving test as one task. Encouraging students to actively use what has been learned from the text to solve problems will lead their standards of coherence to be set on learning new knowledge from the text.

With the implementation of instruction for reading to learn, higher educators can train human resources who can develop necessary knowledge through reading and release them into global society. I believe that this practice constitutes one of the most important educational roles that the university will have in the future.

References

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1061–1070. doi: 10.1037/0278-7393.19.5.1061

Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General*, *128*, 186–197. doi: 10.1037/0096-3445.128.2.186

Barry, S., & Lazarte, A. A. (1998). Evidence for mental models: How do prior knowledge, syntactic complexity, and reading topic affect inference generation in a recall task for nonnative readers of Spanish? *The Modern Language Journal*, *82*, 176–193. doi: 10.1111/j.1540-4781.1998.tb01191.x

Bensoussan, M. (1990). EFL reading as seen through translation and discourse analysis: Narrative vs. expository texts. *English for Specific Purposes*, *9*, 49–66. doi: 10.1016/0889-4906(90)90028-B

Best, R. M., Floyd, R. G., & McNamara, D. S. (2008). Differential competencies contributing to children's comprehension of narrative and expository texts. *Reading Psychology*, *29*, 137–164. doi: 10.1080/02702710801963951

Best, R. M., Rowe, M., Ozuru, Y., & McNamara, D. S. (2005). Deep-level comprehension of science texts: The role of the reader and the text. *Topics in Language Disorders*, *25*, 65–83.

Britton, K. B. (1994). Understanding expository text: Building mental structures to induce insights. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 641-674). San Diego: Academic Press.

Britton, B. K., & Gülgöz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology*, *83*, 329–345. doi: 10.1037/0022-0663.83.3.329

Caldwell, J., & Leslie, L. (2010). Thinking aloud in expository text: Processes and outcomes. *Journal of Literacy Research*, *42*, 308–340. doi: 10.1080/1086296X.2010.504419

Carlson, S. E., van den Broek, P., McMaster, K., Rapp, D. N., Bohn-Gettler, C. M., Kendeou, P., & White, M. J. (2014). Effects of comprehension skill on inference generation during reading. *International Journal of Disability, Development and Education*, *61*, 258–274. doi: 1034912X.2014.934004

Centeno-Cortés, B., & Jiménez Jiménez, A. F. (2004). Problem-solving tasks in a foreign language: The importance of the L1 in private verbal thinking. *International Journal of Applied Linguistics*, *14*, 7–35. doi: 10.1111/j.1473-4192.2004.00052.x

Gerrig, R. J. (2005). The scope of memory-based processing. *Discourse Processes*, *39*, 225–242. doi: 10.1080/0163853X.2005.9651681

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in Instructional Psychology* (pp. 161–238). Mahwah, NJ: Lawrence Erlbaum Associates.

Coté, N., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse Processes*, *25*, 1–53. doi: 10.1080/01638539809545019

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, UK: Cambridge University Press.

Gilliam, S., Magliano, J. P., Millis, K. K., Levinstein, I., & Boonthum, C. (2007).

Assessing the format of the presentation of text in developing a Reading Strategy Assessment Tool (R-SAT). *Behavior Research Methods*, *39*, 199–204. doi: 10.3758/BF03193148

Golding, J. M., Millis, K. K., M., Hauselt, J., & Sego, S. A. (1995). The effect of connectives and causal relatedness on text comprehension. In R. Lorch & E. O'Brien (Eds.), *Sources of coherence in text comprehension* (pp. 127–145). Hillsdale, NJ: Erlbaum.

Goldman, S. R. (1997). Learning from text: Reflections on the past and suggestions for the future. *Discourse Processes*, *23*, 357–398. doi: 10.1080/01638539709544997

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York: Cambridge University Press.

Graesser, A. C., & Bertus, E. L. (1998). The construction of causal inferences while reading expository texts on science and technology. *Scientific Studies of Reading*, *2*, 247–269. doi: 10.1207/s1532799xssr0203_4

Graesser, A. C., Bertus, E. L., & Magliano, J. P. (1995). Inference generation during the comprehension of narrative text. In R. Lorch & E. O'Brien (Eds.), *Sources of coherence in text comprehension* (pp. 295–320). Hillsdale, NJ: Erlbaum.

Graesser, A. C., & Clark, L. F. (1985). *Structures and procedures of implicit knowledge*. Norwood, NJ: Ablex.

Graesser, A. C., & Hemphill, D. (1991). Question answering in the context of scientific mechanisms. *Journal of Memory and Language*, *30*, 186–209. doi: 10.1016/0749-596X(91)90003-3

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, *36*, 193–202. doi: 10.3758/BF03195564

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371–395. doi:10.1037/0033-295X.101.3.371

Halldorson, M., & Singer, M. (2002). Inference processes: Integrating relevant knowledge and text information. *Discourse Processes*, 34, 145–161. doi: 10.1207/S15326950DP3402_2

Horiba, Y. (1996). Comprehension processes in L2 reading: Language competence, textual coherence, and inferences. *Studies in Second Language Acquisition*, *18*, 433–473. doi: 10.1017/S0272263100015370

Horiba, Y. (2000). Reader control in reading: Effects of language competence, text type, and task. *Discourse Processes*, *29*, 223–267. doi:10.1207/S15326950dp2903_3

Horiba, Y. (2013). Task-induced strategic processing in L2 text comprehension. *Reading in a Foreign Language*, 25, 98–125.

Horiba, Y., van den Broek, P., & Fletcher, C. R. (1993). Second Language Readers' Memory for Narrative Texts: Evidence for Structure-Preserving Top-Down Processing. *Language Learning*, *43*, 345–372. doi: 10.1111/j.1467-1770.1993.tb00618.x

Hosoda, M. (2014). Generating causal bridging inferences in EFL expository reading: Combining on-line and off-line processing. *ARELE (Annual Review of English Language Education in Japan)*, *25*, 239–254.

Ishikawa, S., Uemura, T., Kaneda, M., Shimizu, S., Sugimori, N., Tono, Y., . . . Murata, M. (2003). *JACET 8000: JACET list of 8000 basic words*. Tokyo, Japan: JACET.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, *99*, 122–149. doi: 10.1037/0033-295X.99.1.122

Kato, D. (2014). Effects of working memory capacity on on-line goal inference generation in EFL Reading, *ARELE (Annual Review of English Language Education in Japan)*, *25*, 223–237.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*, 294–303. doi: 10.1037/0003-066X.49.4.294

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY: Cambridge University Press.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*, 363–394. doi: 10.1037/0033-295X.85.5.363

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and language*, *29*, 133–159. doi: 10.1016/0749-596X(90)90069-C

Koda, K. (2005). *Insignts into second language reading: A cross-linguistic approach*. New York, NY: Cambridge University Press.

Langston M., & Trabasso T. (1999). Modeling causal integration and availability of information during comprehension of narrative texts. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 29–69). Mahwah, NJ: Lawrence Erlbaum Associates.

León, J. A., & Peñalba, G. E. (2002). Understanding causality and temporal sequence in scientific discourse. In J. Otero, J. A. León & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 155–178). Mahwah, NJ: Erlbaum.

Li, M., & Kirby, J. R. (2014). Unexpected poor comprehenders among adolescent ESL students. *Scientific Studies of Reading*, *18*, 75–93. doi: 10.1080/10888438.2013.775130

Linderholm, T., Everson, M. G., van den Broek, P., Mischinski, M., Crittenden, A., &

Samuels, J. (2000). Effects of causal text revisions on more- and less-skilled readers' comprehension of easy and difficult texts. *Cognition and Instruction*, *18*, 525–556. doi: 10.1207/S1532690XCI1804_4

Linderholm, T., & van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology*, *94*, 778–784. doi: 10.1037/0022-0663.94.4.778

Linderholm, T., Virtue, S., Tzeng, Y., & van den Broek, P. (2004). Fluctuations in the availability of information during reading: Capturing cognitive processes using the landscape model. *Discourse Processes*, *37*, 165–186. doi: 10.1207/s15326950dp3702_5

Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction*, *21*, 251–283. doi: 10.1207/S1532690XCI2103_02

Magliano, J. P., Millis, K. K., Levinstein, I., & Boonthum, C. (2011). Assessing comprehension during reading with the Reading Strategy Assessment Tool (RSAT). *Metacognition and Learning*, *6*, 131–154. doi: 10.1007/s11409-010-9064-2

Magliano, J. P., Millis, K. K., the RSAT Development Team, Levinstein, I., & Boonthum, C. (2011). Assessing comprehension during reading with the Reading Strategy Assessment Tool (RSAT). *Metacognition and Learning*, *6*, 131–154. doi: 10.1007/s11409-010-9064-2

Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999). Strategic processing during comprehension. *Journal of Educational Psychology*, *91*, 615–629. doi: 10.1037/0022-0663.91.4.615

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, *9*, 111–151. doi: 10.1016/0010-0285(77)90006-8

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology*, *93*, 377–389. doi: 10.1037/0022-0663.93.2.377

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. *Psychology of Learning and Motivation*, *51*, 297–384. doi: 10.1016/S0079-7421(09)51009-2

McCrudden, M. T., Magliano, J. P., & Schraw, G. (2011). *Text relevance and learning from text*. Greenwich, CT: Information Age Publishing.

McCrudden, M. T., Schraw, G., & Lehman, S. (2009). The use of adjunct displays to facilitate comprehension of causal relationships in expository text. *Instructional Science*, *37*, 65–86. doi: 10.1007/s11251-007-9036-3

McGraw-Hill. (1992). *McGraw-Hill encyclopedia of science and technology 7th edition*. New York, NY: McGraw-Hill.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, *99*, 440–466. doi: 10.1037/0033-295X.99.3.440

McMaster, K. L., van den Broek, P., Espin, C. A., Pinto, V., Janda, B., Lam, E., ...van Boekel, M. (2015). Developing a reading comprehension intervention: Translating cognitive theory to educational practice. *Contemporary Educational Psychology*, *40*, 28–40. doi: 10.1016/j.cedpsych.2014.04.001

McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes*, *38*, 1–30. doi: 10.1207/s15326950dp3801_1

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, *14*, 1–43. doi: 10.1207/s1532690xci1401_1

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*, *22*, 247–288. doi: 10.1080/01638539609544975

McNamara, D. S., & O'Reilly, T. (2009). Theories of comprehension skill: Knowledge and strategies versus capacity and suppression. In A. M. Columbus (Ed.), *Advances in psychology research* (pp. 113–136). New York, NY: Nova Science Publishers.

Millis, K. K., & Barker, G. P. (1996). Answering questions about expository texts. *Discourse Processes*, *21*, 57–84. doi: 10.1080/01638539609544949

Millis, K. K., Magliano, J., & Todaro, S. (2006). Measuring discourse-level processes with verbal protocols and latent semantic analysis. *Scientific Studies of Reading*, *10*, 225–240. doi: 10.1207/s1532799xssr1003_2

Millis, K. K., Simon, S., & Tenbroek, N. S. (1998). Resource allocation during the rereading of scientific texts. *Memory & Cognition*, *26*, 232–246. doi: 10.3758/BF03201136

Morishima, Y. (2013). Allocation of Limited Cognitive Resources During Text Comprehension in a Second Language. *Discourse Processes*, *50*, 577–597. doi: 10.1080/0163853X.2013.846964

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 519–533. doi: 10.1016/S0022-5371(77)80016-9

Mulder, G. (2008). *Understanding causal coherence relations*. Utrecht, The Netherlands: Landelijke Onderzoeksschool Taalwetenschap (LOT).

Mulder, G., & Sanders, T. J. (2012). Causal coherence relations and levels of discourse representation. *Discourse Processes*, *49*, 501–522. doi: 10.1080/0163853X.2012.692655

Muramoto, T. (2000). *Daini gengo no bunsho rikai katei ni oyobosu shujukudo no eikyo* [The effects of second-language proficiency on text comprehension]. *Science of Reading*, *44*, 43–50.

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language*, *26*, 453–465. doi: 10.1016/0749-596X(87)90101-X

Noordman, L. G. M., & Vonk, W. (1992). Readers' knowledge and the control of inferences in reading. *Language and Cognitive Processes*, *7*, 373–391. doi: 10.1080/01690969208409392

Noordman, L. G. M., & Vonk, W. (1997). The different functions of a conjunction in constructing a representation of the discourse. In J. Costermans & M. Fayol (Eds.), *Processing interclausal relationships: Studies in the production and comprehension of text* (pp. 75–93). Mahwah, NJ: Erlbaum.

Noordman, L. G. M., & Vonk, W. (1998). Memory-based processing in understanding causal information. *Discourse Processes*, *26*, 191–212. doi: 10.1080/01638539809545044

Noordman, L. G. M., Vonk, W., & Kempff, H. J. (1992). Causal inferences during the reading of expository texts. *Journal of Memory and Language*, *13*, 573–590. doi: 0749-596X(92)90029-W

O'Brien, E. J., Rizzella, M. L., Albrecht, J. E., & Halleran, J. G. (1998). Updating a situation model: A memory-based text processing view. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1200–1210. doi:10.1037/0278-7393.24.5.1200

Obunsha. (2005a). *Eiken jyun ichi kyu zen mondai-syu* [The collection of previous questions for the pre-first grade EIKEN test]. Tokyo: Author.

Obunsha. (2005b). *Eiken ni kyu zen mondai-syu* [The collection of previous questions for the second grade EIKEN test]. Tokyo: Author.

O'Reilly, T., & McNamara, D. S. (2007). Reversing the reverse cohesion effect: Good texts can be better for strategic, high-knowledge readers. *Discourse Processes*, *43*, 121–152. doi: 10.1080/01638530709336895

Otero, J., Caldeira, H., & Gomes, C. J. (2004). The influence of the length of causal chains on question asking and on the comprehensibility of scientific texts. *Contemporary Educational Psychology*, *29*, 50–62. doi:10.1016/S0361-476X(03)00018-3

Ozuru, Y., Briner, S., Best, R., & McNamara, D. S. (2010) Contributions of self-explanation to comprehension of high- and low-cohesion texts. *Discourse Processes*, 47, 641–667.

Radvansky, G. A., Tamplin, A. K., Armendarez, J., & Thompson, A. N. (2014). Different kinds of causality in event cognition. *Discourse Processes*, *51*, 601–618. doi: 10.1080/0163853X.2014.903366

Reder, L. M., & Anderson, J. R. (1980). A partial resolution of the paradox of interference: The role of integrating knowledge. *Cognitive Psychology*, *12*, 447–472. doi: 10.1016/0010-0285(80)90016-X

Reder, L. M., & Ross, B. H. (1983). Integrated knowledge in different tasks: The role of retrieval strategy on fan effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 55–72. doi: 10.1037/0278-7393.9.1.55

Robinson, P., & Ross, S. (1996). The development of task-based assessment in English for academic purposes programs. *Applied Linguistics*, *17*, 455–476. doi: applin/17.4.455

Singer, M., & Gagnon, N. (1999). Detecting causal inconsistencies in scientific text. In S. R. Goldman, A. C., Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 179–194). Hillsdale, NJ: Laurence Erlbaum Associates.

Singer, M., & Halldorson, M. (1996). Constructing and validating motive bridging inferences. *Cognitive Psychology*, *30*, 1–38. doi: 10.1006/cogp.1996.0001

Singer, M., Halldorson, M., Lear, J. C., & Andrusiak, P. (1992). Validation of causal bridging inferences in discourse understanding. *Journal of Memory and Language*, 31, 507–524. doi: 10.1016/0749-596X(92)90026-T

Singer, M., Harkness, D., & Stewart, S. T. (1997). Constructing inferences in expository text comprehension. *Discourse Processes*, *24*, 199–228. doi: 10.1080/01638539709545013

Singer, M., & O'Connell, G. (2003). Robust inference processes in the comprehension of expository text. *European Journal of Cognitive Psychology*, *15*, 607–631. doi: 10.1080/095414400340000079

Suh, S. Y., & Trabasso, T. (1993). Inferences during reading: Converging evidence from discourse analysis, talk-aloud protocols, and recognition priming. *Journal of Memory and Language*, *32*, 279–300. doi: 10.1006/jmla.1993.1015

Tannenbaum, R. J., & Wylie, E. C. (2008). *Linking English language test scores onto the Common European Framework of Reference: An application of standard-setting methodology* (ETS Research Rep. No. RR-08-34; TOEFL iBT Research Rep. No. TOEFLiBT-06). Princeton, NJ: Educational Testing Service. Retrieved from https://www.ets.org/Media/Research/pdf/RR-08-34.pdf.

Tannenbaum R. J., & Wylie, E. C. (2013). Mapping TOEIC and TOEIC Bridge Test Scores to the Common European Framework of Reference. In D. E. Powers

(Ed.), *The Research Foundation for the TOEIC Tests: A compendium of studies* (Vol. 2, pp. 6.1–6.10). Princeton, NJ: Educational Testing Service. Retrieved from https://www.ets.org/Media/Research/pdf/TC2-06.pdf

Trabasso, T., & Magliano, J. P. (1996). Conscious understanding during comprehension. *Discourse Processes*, *21*, 255–287. doi: 10.1080/01638539609544959

Trabasso, T., & Suh, S. Y. (1993). Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. *Discourse Processes*, *16*, 3–34. doi: 10.1080/01638539309544827

Trabasso, T., Suh, S. Y., Payton, P., & Jain, R. (1995). Explanatory inferences and other strategies during comprehension and their effects on recall. In R. Lorch & E. O'Brien (Eds.), *Sources of coherence in text comprehension* (pp. 219–240). Hillsdale, NJ: Erlbaum

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, *24*, 612–630. doi:0749596X8590049X

Trabasso, T., van den Broek, P., & Suh, S. Y. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes*, *12*, 1–25. doi:10.1080/01638538909544717

Ushiro, Y., Hamada, A., Kimura, Y., Nahatame, S., Hosoda, M., Kato, D., & Watanabe, Y. (2015). Building causally coherent mental representations and learning from expository texts among Japanese EFL readers. *JACET Journal*, *59*, 131–150.

Ushiro, Y., Nahatame, S., Hasegawa, Y., Kimura, Y., Hamada, A., & Tanaka, N. (2014). Narrative characters' goals and EFL readers' text comprehension: Focusing on goal explicitness. *ARELE*, *25*, 1–16.

Ushiro, Y., Shimizu, H., Kai, A., Nakagawa, C., Takaki, S., Kobayashi, M., Satake, N., & Takano, D. (2010). Effects of causal networks on on-line and off-line narrative comprehension among Japanese EFL readers. *JACET Journal*, *51*, 39–53.

van den Broek, P. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science*, *328*, 453–456. doi:10.1126/science.1182594

van den Broek, P., Lorch, R. F., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition*, *29*, 1081–1087. doi: 10.3758/BF03206376

van den Broek, P., Virtue, S., Everson, M. G., Tzeng, Y., & Sung, Y. C. (2002). Comprehension and memory of science texts: Inferential processes and the construction of a mental representation. In J. Otero, J. A. León & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 131–154). Mahwah, NJ: Erlbaum.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic Press.

Varnhagen, C. K. (1991). Text relations and recall for expository prose. *Discourse Processes*, *14*, 399–422. doi:10.1080/01638539109544794

Vidal-Abarca, E., Martínez, G., & Gilabert, R. (2000). Two procedures to improve instructional text: Effects on memory and learning. *Journal of Educational Psychology*, *92*, 107–116. doi:10.1037/0022-0663.92.1.107

Voss, J. F., & Silfies, L. N. (1996). Learning from history text: The interaction of knowledge and comprehension skill with text structure. *Cognition and Instruction*, *14*, 45–68. doi: 10.1207/s1532690xci1401_2

Weaver, C. A. III., & Kintsch, W. (1991). Expository text. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research volume II* (pp.

230–245). Mahwah, NJ. Lawrence Erlbaum Associates.

Wiley, J., & Myers, J. L. (2003). Availability and accessibility of information and causal inferences from expository text. *Discourse Processes*, *36*, 109–129. doi: 10.1207/S15326950DP3602_2

Wittwer, J., & Ihme, N. (2014). Reading skill moderates the impact of semantic similarity and causal specificity on the coherence of explanations. *Discourse Processes*, *51*, 143–166. doi: 10.1080/0163853X.2013.855577

Yoshida, M. (2003). Working memory capacity and the use of inference in L2 reading. *JACET Bulliten*, *40*, 57–82.

Zwaan, R. A., & Brown, C. M. (1996). The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Processes*, *21*, 289–327. doi: 10.1080/01638539609544960

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162–185. doi: 10.1037/0033-2909.123.2.162

Appendices

Appendix A: The Causal Question

英文の内容を順序立てて説明する課題を行っていただきます。
注意点 原因と結果を、論理的に正しい順番に最初から並べて、可能な限り詳しく説明してください。

(例題) 地震が起こるのはなぜですか。その原因を最初から順序立てて，可能な限り詳しく説明してください。

地球のプレートが移動し、海のプレートと陸のプレートがぶつかり合うと陸のプレー
ト側が巻き込まれ、巻き込まれたプレートの端が反発することで地震が発生する。

Q1 無重力空間にいると、心臓が小さくなるのはなぜですか？ 原因と結果を出来るだけあげながら、可能な限り詳しく説明してください。
無重力空間にいると、
心臓が小さくなる。

Appendix B: The Problem-Solving Test

文章に書かれていた内容を、文章とは別の状況や、問題の解決へ応用する課題を行っていただきます。

注意点

(1) 文章から学んだことを最大限に使って、それぞれを出来る限り詳しく説明してください。

(2) 結果を書くだけではなく、なぜそうなるのかを具体的に理由付けてください。

(例題) 火山がある場所では、地震がよく起こります。これはなぜだと考えられますか？

火山の周りの地下はマグマの通り道になっており、それが移動することで周囲の岩盤が割れて、震動が起こるから。

Q1	宇宙で肺や心臓が文章にあったような命令を出しても，体の他の部位 (内臓など) がその命令に従わなかった場合，体の水分レベルはどうなると考えられますか？※宇宙での話しです。
Q2	心臓の動きが弱っているときは，水分を摂り過ぎないよう注意が必要です。この理由を，文章中の心臓と体内の水分量の関係から学んだことを基に説明してください。※地上での話しです。
Q3	体が，体内の水分の変化に気づけなくなる病気にかかったと想像してください。この状態で宇宙に行っても，心臓の大きさは地上とあまり変わらないと考えられます。この理由を，文章から学んだことを基に説明してください。※宇宙での話しです。
Q4	宇宙飛行士の心臓が小さくなることを防ぐには、宇宙ステーションにはどのような設備や機能が必要だと考えられますか？

Appendix C: Experimental Texts, Inference Questions, and Detail Questions in Experiment 4

IQ = inference question, DQ = detail question, Italicized and underlined sentences are direct and target sentences, respectively.

<u>High-Familiarity Texts</u>

The cacao tree is a small tropical tree whose seeds are used to make chocolate.

In general, direct sunlight damages the cacao's growth.

<u>Farmers prefer woody areas for cacao cultivation because it receives less sunlight</u>.

Traditionally, people farmed the cacao in forests.

IQ: カカオの栽培に直射日光は適さない

DQ: カカオは森で栽培されていた

At the end of a star's life, its gravity becomes extremely strong.

Stars that have strong gravity live shorter.

<u>Big stars have much shorter lives than small stars because big stars' gravity is far stronger</u>.

Most black holes are considered to be the result of the death of such large stars.

IQ: 重力の強い星は短命だ

DQ: 大きな星の死はブラックホールの原因になる

Miso is a traditional Japanese seasoning, made from soybeans, sea salt, and rice.

Among them, salt has the effect of raising body temperature.

<u>Miso is sometimes recommended for keeping the body warm because it contains a rich</u>

amount of salt.

Nowadays, food companies produce miso in large quantities and the home-made style has become rare.

IQ: 塩分は体を温かく保つ

DQ: 手作りのミソは今でもよくみられる

Spinach (ほうれん草) is a highly popular green vegetable that can be eaten either raw or cooked.

Spinach can be damaged by diseases in conditions of high temperature.

In farming spinach, hot areas are usually avoided because diseases can spoil harvest.

In terms of nutrition, spinach is extremely rich in iron, vitamin A, and vitamin C.

IQ: ほうれんそうは暑いと病気にかかる

DQ: ほうれんそうは野菜の中でも人気がない

Climate change is predicted to have a serious impact on worldwide agriculture.

Taking potatoes for example, hot climates are not suitable for their cultivation.

Researchers predict reductions in potato production because the global climate is rapidly warming.

Owing to this, researchers also expected that the banana might replace potatoes in some developing countries.

IQ: 暑いとポテトの生産量が低下する

DQ: 気候変動の影響は発展途上国に限られる

The tiger is the largest cat species, reaching a total body length of three meters.

Nocturnal (夜行性の) animals have very poor color vision.

Tigers cannot distinguish colors because they are active at night.

Over 95% of tigers have been lost in the last 100 years, making them an endangered species.

IQ: 夜行性の動物は色を区別できない

DQ: トラは夜行性の動物だ

The melon is a widely eaten fruit that takes two or three months to ripen (熟成する).

The melon makes other fruits ripe faster when it is placed nearby.

Farmers sometimes store the melon with other fruits because they want such fruits to ripen faster.

The price of the melons greatly varies depending on species and region.

IQ: メロンは他の果物の熟成を早める

DQ: メロンは熟成するのに半年以上かかる

Macadamia nuts are the fruit of a tropical tree.

High quality nuts float in water.

In Hawaii, the nuts can be seen to be of high quality because they float in water.

Before packed, they are roasted (炒られる) and salted.

IQ: 質の高いマカダミアナッツは水に浮く

DQ: マカダミアナッツは梱包前に炒られる

Color affects people's various feelings such as imagination and relief.

For example, violet is known to develop one's creativity.

Purple color is often favored by artists because they wish to increase their imagination.

On the other hand, some hospitals make use of green color to relieve patients.

IQ: 紫色は想像力を高める

DQ: アーティストは紫色を好む

Skiers of different abilities need different equipment, such as compact or large one.

It is easier to turn on shorter, more compact skis.

<u>Beginners are frequently advised to use compact skis because they usually have difficulty in changing direction</u>.

Once they can control their movements, they can quickly advance in skill.

IQ: コンパクトなスキー板は曲がりやすい

DQ: スキーの初心者は上達に時間がかかる

Aerobics (有酸素運動) are forms of physical exercises, effective for burning fat, such as jogging, and swimming.

However, research has shown that aerobics can lead to muscle loss.

<u>Some bodybuilders avoid aerobics because the loss of muscle size can be a critical problem for them</u>.

Those who want to stay healthy often prefer the combination of aerobics and weight training.

IQ: 有酸素運動は筋肉量を失わせる

DQ: 有酸素運動はウェイトトレーニングと組み合わせられる

Frost (霜) damage has been a major concern among farmers in fruit-producing regions.

Big winds reduce frost formation.

<u>In some countries, farmers use a helicopter to prevent frost because it can easily create strong winds</u>.

The best protection for avoiding frost is locating the fruit planting on a proper site.

IQ: 強風は霜の発生を防ぐ

DQ: 霜の被害は猟師にとって最大の懸念事項だ

Low-Familiarity Texts

Morning Glory clouds are very rare roll clouds that can be observed in Burketown, Australia.

Studies have shown that one cause of the Morning Glory clouds is low pressure.

The Morning Glory clouds are most often observed in August because low pressure is to the south of Burketown.

The formations of the Morning Glory clouds have been noticed since ancient times.

IQ: 低気圧はモーニンググローリーを起こす

DQ: モーニンググローリーは春に頻繁に発生する

Underwater (水中の) cables are used to transmit power (電力を送る) to distant islands.

Cooled cables have greater power capability.

The power capacity of underwater cables is increased because such cables are naturally cooled.

However, underwater cables are sometimes damaged by ships.

IQ: 冷却されたケーブルは電力容量が大きい

DQ: 水中ケーブルは船を傷つける

Tsetse flies (ツェツェバエ) are harmful insects living in tropical Africa.

Tsetse flies cause sleeping sickness.

The control against tsetse flies is important because sleeping sickness is extremely dangerous.

Some organic treatments are effective, especially when conducted early in the disease.

IQ: ツェツェバエは眠り病を引き起こす

DQ: 眠り病は軽い病気だ

Barges (荷揚げ船) are large boats that are used to transport heavy goods.

Rivers must be nine feet deep to support barges.

The James River can float barges because it is over nine feet in depth.

One barge carries up to five times its weight in freight (貨物).

IQ: 荷揚げ船を浮かべるには 9 フィートの深さがいる

DQ: 荷揚げ船は人を運ぶために利用される

Multiple personality disorder (統合失調症) is a rare mental disorder, characterized by sconfusion in thinking.

Drugs for multiple personality disorder can cause disturbances of movement.

Some patients with multiple personality disorder experience a lack of balance because they have been treated with drugs.

Drug treatment is only partially successful in that normal functioning is not completely restored by drugs.

IQ: 統合失調症の薬はバランスの欠如を起こす

DQ: 統合失調症は薬物治療だけでは完治できない

In ancient times, the sizes of animals and plants were much larger than today.

The animals get bigger when the atmosphere contains a larger amount of oxygen.

The ancient animals grew to such large sizes because the air oxygen level was very high.

In addition, climatic and genetic factors interactively influence specific animal body sizes.

IQ: 酸素濃度が高いと動物は大きくなる

DQ: 遺伝は動物の大きさに独立した影響を与える

Millets (アワ) are a group of grains that have a longer history of cultivation than rice.

Low humidity areas are suitable for the large production of millets'.

One of the millets' main producing areas is Africa because there is low humidity.

Millets are available in markets throughout the year.

IQ: アワの栽培には低い湿度が適している

DQ: アワは米よりも新しい穀物だ

La Nina (ラニーニャ現象) consists of the cooling of surface ocean waters in the Pacific Ocean along the western coast of South America.

One effect of La Nina is stronger easterly trade winds.

1988 saw unusually strong easterly trade winds because the world experienced the largest La Nina.

This in turn caused droughts in United States, decreasing corn yields by 30 per cent from the previous year.

IQ: ラニーニャ現象は東貿易風を強める

DQ: 1988 年の東貿易風はとても強かった

A photoconductor (導光体) is a part of copy machines to improve the flow of electricity.

In making photoconductors, manufacturers often use silicon, which has positive electricity.

<u>Photoconductors in commercial copy machines are positively charged because they are composed of silicon parts</u>.

Today, photoconductors are a vital component of many appliances.

IQ: シリコンは正の電気を帯びる

DQ: 導光体は多くの製品で使われている

Ijen volcano (イジェン火山) in Indonesia is known to be a sulfur (硫黄) mining operation site.

Sulfur is known to produce blue fire.

<u>Flames from Ijen volcano gleam in mysterious blue because the magma is filled with molten sulfur</u>.

The mining involves inferior working conditions, leaving workers with a life expectancy of just 30 years.

IQ. 硫黄のマグマは青く燃える

DQ: イジェン火山の労働環境は改善されている

Locusts plague (イナゴ被害) is a devastating natural disaster, causing the loss of human lives as well as crops.

Research has shown that a drought can lead to a locust plague.

<u>China suffered severe damage from locust plagues because a high frequency of droughts hit the country</u>.

Even today, the disaster threatens lots of countries, many of which do not have sufficient training or funds.

IQ: 干ばつはイナゴ被害を起こす

DQ: イナゴ被害への対策が不十分な国は少なくない

Infrared (赤外線) is a light wave that is just beyond the visible range.

Studies reported that infrared is produced from stars whose surface temperature is less than 4000℃.

Betelgeuse produces most of its energy in infrared because its surface temperature does not reach 4000 ℃.

Examining infrared rays is one method to measure distances to stars.

IQ: 4000℃未満の星は赤外線を発する

DQ: ベテルギウスの表面温度は 5000℃を超える

Appendix D: An Experimental Text in the Explicit and Implicit Conditions in Experiments 5 and 6

<table>
<tr><th>Explicit condition</th><th>Implicit Condition</th></tr>
<tr><td colspan="2">When people first considered space travel, they did not know how the zero gravity of space would affect humans. In fact, the human body is a complex system that automatically responds to the lack of gravity.</td></tr>
<tr><td>While in space, the body is not affected by gravity. Therefore, blood and water do not travel to the lower parts of the body, especially the legs. Instead, the blood and water within the body move to the upper body. Because the blood and water travel to the upper parts of the body, the body feels like the chest and head are filled with blood and water. Because of this, the heart and lungs send messages that the amount of blood and water in the upper part of the body must be reduced. As a result, space travelers do not feel thirsty, and therefore, space travelers drink less water. As body water is eliminated, their body water levels become lower than normal. When the amounts of blood and water decrease, it becomes more difficult for the human body to work normally. In addition, the decreased body water makes the heart pump less blood than normal. Therefore, the heart does not need to work as hard as it does on Earth. As a result, the heart becomes smaller.</td><td>While in space, the body is not affected by gravity. Blood and water do not travel to the lower parts of the body, especially the legs. They move to the upper body. The body feels like the chest and head are filled with blood and water. The heart and lungs send messages that the amount of blood and water in the upper part of the body must be reduced. Space travelers do not feel thirsty, and they drink less water. Body water is eliminated, and the body water levels become lower than normal. It becomes more difficult for the human body to work normally. The heart pumps less blood than normal. It does not need to work as hard as it does on Earth. The heart becomes smaller.</td></tr>
<tr><td colspan="2">Studying the effects of space travel on humans can help us better understand many illnesses, such as high blood pressure and other heart problems.</td></tr>
</table>

Note. Bolded text = connectives and other linguistic cues added in the explicit text; underlined text = pronoun and anaphoric expressions replaced with nouns in the explicit text; double-underlined text = explanatory information added in the explicit text.

Appendix E: An Experimental Text (in the Explicit Condition) and the Number of TCCs and ECCs in Experiments 5 and 6

Bold text = target statements for thinking aloud in Experiment 6. Statements 9 and 18's causal relations were not computed because they were not included in the implicit-condition text.

	Statement	The number of TCCs	TCC Group	The number of ECCs	ECC Group
1	When people first considered space travel	0	Low	0	Low
2	they did not know	0	Low	0	Low
3	**how the zero gravity of space would affect humans.**	14	High	0	Low
4	The human body is a complex system	1	Low	0	Low
5	**that automatically responds to the lack of gravity**	12	Middle	1	Low
6	**While in space, the body is not affected by gravity.**	16	High	0	Low
7	Therefore, blood and water do not travel to the lower parts of the body, especially the legs.	15	High	2	Middle
8	**Instead, the blood and water within the body move to the upper body.**	15	High	3	Middle
9	Because the blood and water travel to the upper parts of the body,				
10	the body feels	4	Low	4	Middle
11	**like the chest and head are filled with blood and water.**	14	High	4	Middle
12	Because of this, the heart and lungs send messages	12	Middle	4	Middle
13	that the amount of blood and water in the upper part of the body must be reduced.	12	Middle	6	High
14	As a result, space travelers do not feel thirsty,	15	High	8	High
15	**and therefore, space travelers drink less water.**	15	High	9	High
16	As body water is eliminated,	12	Middle	10	High
17	**their body water levels become lower than normal.**	15	High	11	High
18	When the amounts of blood and water decrease,				
19	it becomes more difficult for the human body	11	Middle	11	High

20	**to work normally.**	3	High	3	Middle
21	In addition, the decreased body water makes the heart pumps less blood than normal.	13	Middle	11	High
22	Therefore, the heart does not need to work	13	Middle	13	High
23	as hard as it does on Earth.	2	Low	2	Middle
24	**As a result, the heart becomes smaller.**	15	High	15	High