Text Cohesion in EFL Reading:
Evidence From Text Analyses and Cohesion Manipulation Studies on Japanese Learners

A Dissertation Submitted to the University of Tsukuba in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Linguistics

Tomoko OGISO

2021

Abstract

Text Cohesion in EFL Reading: Evidence From Text Analyses and Cohesion Manipulation Studies on Japanese Learners

By

Tomoko OGISO

Reading comprehension is prevalent throughout our lives. We are exposed to text-based information through books, newspapers, websites, and social networking sites. In today's society, where information and communication technologies are highly developed, and countries worldwide communicate through the internet, we can immediately access texts written by native speakers and writers from diverse positions via various media. In such circumstances, in daily life, work, and academic contexts, it is important to accurately understand information from texts in both native and foreign languages, connect that information, and read it in a defensible and consistent manner. Particularly in foreign language learning, where students are unexposed to sufficient language input, the development of reading comprehension skills beyond a surface understanding of the text, including a rich construction of text representation, is sometimes challenging.

Previous research on reading comprehension has examined native and foreign language reading processing and text- and reader-related factors that affect reading comprehension. Reading comprehension is achieved through an interaction between the text and reader. Additionally, a coherent understanding of a text requires an appropriate connection and explicit presentation of the text's words and phrases and the construction of a situational model through the reader's inference generation (e.g., Kintsch, 1998).

However, not all texts are appropriately presented to readers. Texts that lack sufficient explanation or explicitness are burdensome to readers (Beck et al., 1991). Moreover, it is difficult for immature learners, especially learners of English as a foreign language (EFL), to infer non-explicit information and causal relationships in a text if the reader does not have the reasoning ability or cognitive resources to adequately understand (e.g., Horiba, 1996).

To cope with these problems, research on reading comprehension has been actively examining the effects of text modification for readers facing difficulties. Previous studies have focused on *cohesion*, which is a textual factor referring to the degree of explicitness of the contents and relations of elements within a text; furthermore, existing literature has investigated the effects of manipulating text cohesion on readers' text comprehension when reading their native language (e.g., Linderholm et al., 2000; McNamara et al., 1996; O'Reilly & McNamara, 2007; Vidal-Abarca et al., 2000).

However, very few studies have attempted to uncover the effects of text cohesion on EFL learners (e.g., Horiba, 1996; Hosoda, 2016), who have limited cognitive resources, compared to native speakers. Additionally, these studies have focused on the effects of text cohesion on readers' comprehension but have not fully elucidated EFL learners' processing during reading. It is revealed that readers' processes, such as connecting sentences, finding relations between them, and visualizing information not explicitly described in the text (i.e., *inference generation*), contribute to readers' construction of coherent understanding. Thus, exploring the effects of text cohesion on readers' processes is needed. Clarifying the effects of cohesion on readers' processing, comprehension, and inference generation and applying them toward teaching materials and instructions can contribute to the development of autonomous readers.

Therefore, this study aims to elucidate the current state of cohesion in English

language materials for Japanese EFL learners (Study 1) and examine the effects of manipulating cohesion (high and low) on Japanese EFL learners' reading comprehension from several empirical studies (Study 2). Study 1 comprised two text analyses (Text Analyses 1 and 2), and Study 2 consisted of six empirical studies (Experiments 1, 2-A, 2-B, 3-A, 3-B, and 4).

Text Analyses 1 and 2 examined whether the cohesiveness of textbooks and EIKEN textbooks varies across levels. Coh-Metrix, an automatic text analysis tool, was used to analyze the characteristics of the English text in 20 foreign language textbooks and teaching materials from elementary to high school (Text Analysis 1) and 501 textbooks of past EIKEN examinations (Text Analysis 2) to observe the changes in level of difficulty. The results showed that one of the cohesion indicators (reference cohesion, the proportion of the content words that overlap between adjacent sentences) was lower for both the textbooks and EIKEN as the text's difficulty or grade level increased.

Experiment 1 examined the effects of text cohesion and readers' proficiency in their sensitivity to cohesion. Thirty Japanese undergraduate and graduate students were surveyed to determine their perceived differences in cohesion while reading high- or low-cohesion texts. Specifically, their sensitivity to cohesion was measured using a cohesion judgment task (cf. Helder et al., 2016) for rating the cohesion of the texts on a six-point scale. The results showed that the rated cohesiveness values of different cohesive text readings were not significantly different between the high- and low-cohesion texts. The results also showed no significant effect of reader proficiency in readers' sensitivity to the judgment of cohesion. This suggests that learners, regardless of their proficiency level, may not consciously pay attention to the meta-characteristics of the text or be able to discern the difference between high- and low-cohesion texts while consciously or unconsciously paying attention to the text characteristics.

Experiment 2 (2-A and 2-B) examined the effect of high and low text cohesion on readers' comprehension of expository texts on social topics. Thirty-six Japanese university students participated in the experiments (a relatively high-proficiency group was observed in Experiment 2-A and a relatively low-proficiency group in the follow-up Experiment 2-B), and their text comprehension was compared between low- and high-cohesion texts. Readers' comprehension was measured using a written recall task in which participants wrote in their L1 (native language, Japanese) whatever they remembered without referring to the L2 (second language, English) texts. Further, a why-question task asked participants to write answers in their L1 about causal relationships in the L2 text. In the written recall tasks, high-proficiency readers recalled more information in the low-cohesion texts than in the high-cohesion texts (Experiment 2-A), whereas the low-proficiency readers' text comprehension did not differ between the cohesion conditions (Experiment 2-A). Therefore, it was confirmed that the cohesion effect in promoting understanding differed depending on the proficiency level of the readers.

Experiment 3 (3-A and 3-B) investigated the effects of high- and low-cohesion text on readers' comprehension of expository texts on scientific topics. Sixty Japanese university students participated in the experiments (a relatively high-proficiency group was surveyed in Experiment 3-A and a relatively low-proficiency group in the follow-up Experiment 3-B), and their comprehension of scientific explanations was compared between high- and low-cohesion texts. As in Experiment 2, the written recall tasks, executed in L1 (Japanese), in which participants wrote what they remembered and responded to a why-question about cause-and-effect relationships after reading the L2 (English) text, were used to measure the readers' comprehension. The results showed that in the relatively high-proficiency group (Experiment 3-A), there was no significant difference between the readers' text comprehension in the high-cohesion or low-cohesion

conditions. In contrast, the relatively low-proficiency group (Experiment 3-B) reported higher comprehension of the high-cohesion texts than the low-cohesion texts. Alternatively, it was established that low-proficiency readers benefited from high-cohesion, especially in texts regarding scientific topics.

Experiment 4 examined the effects of text cohesion and readers' proficiency in reading English and their processing during reading. Employing the think-aloud method, which asks readers to verbalize their thoughts during reading, think-aloud protocols from 40 Japanese students who participated in Experiment 3 were analyzed in terms of reading processes while reading high- and low-cohesion texts about scientific explanations. Specifically, participants' inference generation during reading was the focus, and the participants' production rates in different cohesive texts were analyzed. Additionally, readers' proficiency levels, measured by a reading proficiency test, and processing tendencies, determined by cluster analysis—lower-level (e.g., word, phrase analyzers) and higher-level (e.g., inference, association producers) processes groups—were considered in the analysis. Considering learners' proficiency level and process-allocation tendencies, the results showed that learners with high proficiency, who performed much higher-level processing, produced more inferences in low-cohesion texts than in high-cohesion texts. In contrast, learners with low proficiency, who engaged more in lower-level processing, generated more inferences when reading the high-cohesion texts than the low-cohesion texts. Therefore, Experiment 4 suggested that readers with different proficiency levels and processing tendencies varied their processing (i.e., inference generation) according to text cohesion. Low-cohesion texts facilitated higher-processing readers' inference generation, while high-cohesion texts helped lower-processing readers' inference generation.

Consequently, Study 1, which analyzed textbooks and large-scale tests for Japanese

learners of English, revealed that the more complex or difficult the English text, the lower the reference cohesion of the text. This indicates that cohesion reflects the difficulty of the text to some extent, as in native language studies. Study 2's results, comprising six empirical experiments, revealed that the degree of text cohesion affects Japanese learners' processing and text comprehension. Moreover, text cohesion has varied effects on EFL learners' comprehension depending on the text genre, readers' proficiency, and readers' process-allocation tendencies during reading.

In conclusion, the present study demonstrates the importance of text cohesion and the interaction between the factors related to cohesion, genre, and EFL reading comprehension. Although there are some limitations to be noted, such as limited sample sizes and methodological problems (i.e., the validity of the comprehension measurements, task burden for participants to produce think-aloud protocols during reading), this study provides new insights into EFL reading and implications for future text development and classroom reading instructions.

Acknowledgements

I would like to express gratitude to all people who assisted me in completing this dissertation. Although it would be difficult to acknowledge every person who helped me here, I would like to express my deepest appreciation to the following people.

First, I would like to express my most profound gratitude to my academic supervisor, Professor Yuji Ushiro, at the University of Tsukuba. He guided my interest in English education and second language reading and has always motivated me to set and purse higher goals. He has continuously given me constructive comments and warm encouragement. I have learned a lot about theories and educational issues in English education and second language learning from his lectures, seminars, and co-research group (Reading Research Group) meetings. Without his guidance and persistent help, I could not have engaged in and completed my research.

Second, I would like to express my appreciation to Professor Hirosada Iwasaki at the University of Tsukuba. He has supported me and provided insightful comments. I learned vocabulary learning in second language and how to write academic theses from his lectures. I really appreciate his consistent encouragement from my master thesis to the Ph.D. dissertation writing.

Assistant Professor Yuko Hijikata at the University of Tsukuba also provided me with incisive comments and essential suggestions from the perspectives of second language reading and theories. She has supported me with the thesis and allowed me to join a co-research project about multiple-document reading. I have also learned the desirable attitude and milestones toward achieving academic success from her.

Next, I would like to extend my appreciation to my external supervisor, Professor Yukie Horiba, at Kanda University of International Studies. She gave me valuable

suggestions in terms of second language acquisition, and her comments helped me revise and improve the earlier version of the dissertation.

Further, I am grateful for the help and support from the Japan Society for the Promotion of Science (JSPS), the EIKEN Foundation of Japan, and the Japan Textbook Research Center. They provided me with research funding and opportunities to conduct text analyses and experiments.

I would also like to express my gratitude to all professors in the Modern Languages and Cultures, Degree Program in Humanities and Social Sciences, the University of Tsukuba for giving me illuminating comments and suggestions for my research.

I am indebted to my senior and junior associates in Professor Ushiro's seminar. Especially, I would like to express my heartfelt appreciation to Dr. Yukino Kimura, Dr. Akira Hamada, Dr. Natsumi Tanaka, Dr. Masaya Hosoda, Dr. Yoshinobu Mori, Dr. Kentaro Suzuki, and Dr. Kozo Kamimura, Mr. Komuro Ryuya for their comments, suggestions, and warm support in the long-term process of completing my doctoral course.

I am immensely thankful to my colleagues, Mr. Yamato Sasaki, Mr. Hideaki Oka, and Mr. Hiroki Maeda, who worked with me during my research and shared both good and hard times during the doctoral course.

Finally, I would like to express profound gratitude to my family and friends for their understanding and irreplaceable support. Without their encouragement, this thesis would not be materialized. Again, I am deeply grateful to every person concerned with this study, including the participants in my experimental studies.

Tomoko OGISO

Table of Contents

List of Tables

Table Page

List of Figures

Figure Page

List of Appendices

Publications

This dissertation covers text analyses and experimental studies reported in the publications listed below:

Ogiso, T. (2018). Effects of text cohesion on inference generation during EFL reading: Evidence from think-aloud protocols. *ARELE: Annual Review of English Language Education in Japan, 29*, 161–176. (Chapter 4)

Ogiso, T. (2019). Effects of text revision on EFL readers' comprehension: Simplified vs. elaborated text. *ARELE: Annual Review of English Language Education in Japan*, *30*, 241–255. (Chapter 4)

小木曽智子. (2018).「英文読解におけるモニタリング能力の測定－一貫性と結束性の比較を通して－」.『EIKEN BULLETIN』, 30 号, 12–26. (Chapters 3 & 4)

小木曽智子. (2020).「小中高の一貫した外国語教科書の作成を目指して－CohMetrix による英文，CEFR-J による語彙の解析をもとに－」.『公益財団法人教科書研究センター「論文集」』. 111–124. (Chapter 3)

Ogiso, T. (2020). Reverse cohesion effect in second-language reading comprehension: A preliminary study. *The 30th Annual Meeting of the Society for Text & Discourse*. Easy Chair Preprint, No.3835. (Chapter 4)

Chapter 1

Introduction

1.1 Background of the Current Research

Reading text is one of the most important ways of communicating and conveying information. In this century, numerous written information is conveyed through electronic networks across international borders for business or politics (e.g., telecommunication and social network services). Conventional written materials are also used for education (e.g., textbooks and homework), academic purposes (e.g., thesis and report), and mass media (e.g., newspapers and magazines). Therefore, reading becomes a powerful daily activity that influences people's lives.

Reading is accomplished through the interaction between the reader and text. Reading is treated as a complex cognitive activity because many factors affect a reader's attitude, processes of reading, and text comprehension (e.g., the purpose of reading, reader's ability, text type). When readers read a text, they are engaged in various cognitive activities, including analyzing words, paraphrasing the sentence, generating inferences relating to text content, and self-monitoring their comprehension. Moreover, text comprehension is the understanding of not only the explicit message stated in the individual words of the text but also the broader message conveyed by a writer. To achieve a successful text comprehension, readers have to construct a meaningful representation of the text in memory based on coherence relations (e.g., van Dijk & Kintsch, 1983). It is assumed that readers encode propositions and relationships and generate inferences from their general knowledge in the process of constructing meaningful representations of the text (e.g., Kintsch, 1988). To maintain a coherent representation in mind, readers must generate inferences; when readers make inferences to construct the global message of the

text, deeper comprehension is achieved (Graesser et al., 1994).

However, inferencing is regarded as an arduous process for readers. Previous studies have indicated the importance of not only lower-level or basic processing skills (e.g., word recognition, syntactic processing, capturing the meaning) but also higher-level processing skills, such as inference skills; they specified that a lack of inference skills might cause difficulty in reading (Kintsch & van Dijk, 1978; Yuill & Oakhill, 1991). The inference skill is a more intellectual process than the lower-level processes where readers understand explicit ideas written in the text. To generate inferences, readers must guess implicit ideas, relationships, and causalities by connecting their knowledge to the text's messages. It is a high-load process for readers because their cognitive resources are limited. Thus, it has been argued that inference skill is one of the characteristics of good readers (Cain & Oakhill, 1999; Oakhill, 1984), and poor readers do not have sufficient ability to infer.

Further, it is especially difficult for readers learning English as a foreign language (EFL) to generate these inferences appropriately while reading a text (Horiba, 1996). When a writer is unskilled or has special intentions (e.g., in detective stories), their message is not always explicitly stated, and cues to connect the relationships of the events in the text are not always provided. In that case, readers have to engage in high-load inferencing to understand the text coherently, and it is harder to comprehend the text, especially for poor readers.

Additionally, the reader's skills (e.g., generating inferences) and text characteristics (e.g., explicitness, readability) are significant factors that can influence readers' text comprehension. To support poor readers' text comprehension, researchers of native language (L1) and second language (L2) reading tried to determine the relationship between the reader- and text-related factors. Some studies have used manipulated text to

enhance its explicitness and cohesiveness and investigated its effects on reading (e.g., Gilabert et al., 2005; Linderholm et al., 2000; Oh, 2001; Vidal-Abarca et al., 2000). These studies showed positive or negative effects of text-related factors on readers' comprehension. However, these studies are mostly for L1 readers, and there is room for investigations for L2 learners. Considering that learners have more difficulties in reading L2 texts compared to native readers, he impact of text modification on L2 learners should be investigated. Additionally, these prior studies focused on readers' comprehension (e.g., reader's recall protocol, the correct number of comprehension questions), not the interaction or processes between readers and texts (e.g., reading process, constructing text representation). Further, there is a lack of research on cohesion judgment, that is, whether readers evaluate the text cohesiveness appropriately. Studies investigating the cohesion of English texts for current EFL learners are rare and lack research with direct implications for the reality of English education in Japan.

1.2 Organization of the Dissertation

Therefore, this study aims to investigate the current state of cohesion in English language materials for Japanese EFL learners (Study 1) and examine the effects of manipulating cohesion (high and low) on Japanese EFL reading with several aspects, such as readers' judgment, reading process, and text comprehension from several empirical studies (Study 2). The present study will be the first step in demonstrating text analysis and comprehensively investigating the effects of text cohesion on Japanese EFL readers. This study's findings have pedagogical implications for researchers, educators, and content developers interested in the interaction between text-reader factors in EFL reading.

Given the background above, this study addressed the following main research questions: (1) How do text characteristics of different-level English materials for

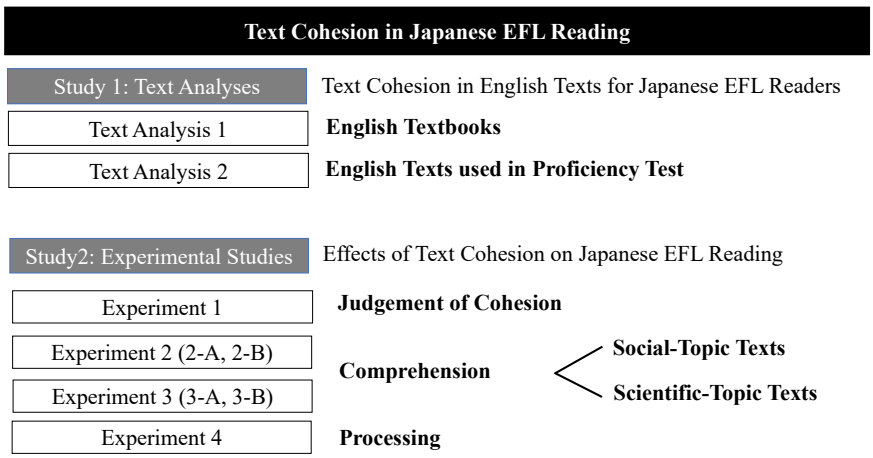
Japanese EFL learners change in terms of text cohesion? and (2) How does manipulating cohesion of English text (i.e., high- and low-cohesion texts) affect Japanese EFL learners' cohesion judgment, text comprehension, and processing during reading?

To answer these questions, this dissertation comprises the following five chapters: Introduction (Chapter 1), Review of Related Literature (Chapter 2), Study 1: Text Analysis on EFL Reading Materials (Chapter 3), Study 2: Effects of Text Cohesion on Japanese EFL Reading (Chapter 4), and General Discussion and Conclusions (Chapter 5).

This thesis includes two major studies (Studies 1 and 2); two text analyses and six experiments were conducted under them. In particular, Text Analyses 1 and 2, which constituted Study 1, aimed at investigating whether the cohesiveness of English textbooks and EIKEN texts differed across the different levels. Study 2, which examined the effects of high- and low-cohesion texts on multiple aspects of Japanese EFL learners' reading (judgment, comprehension, processing), comprised a total of six experiments—four main and two follow-up experiments (see Figure 1.1).

Figure 1.1

Overview of the Two Text Analyses and Six Experimental Studies in the Present Study



Chapter 2 reviews the literature related to the present study. Specifically, theories of reading comprehension and processing (e.g., construction of mental representation in reading, the construction-integration model, inference generation during reading), text cohesion and coherence in reading, effects of text revision on readers' coherence judgment, and comprehension revealed by L1 and L2 experimental studies are reviewed. The findings and limitations are summarized at the end of this chapter.

In Chapter 3, Study 1, the first study, which uncovered the textual features of Japanese EFL reading materials using an automatic textual analyzer, Coh-Metrix (e.g., Crossley et al., 2012; Crossley et al., 2007, Graesser et al., 2004; McNamara et al., 2014), is presented. Study 1 comprised two investigations of text analyses (Text Analyses 1 and 2). For Text Analysis 1, English textbooks for elementary, middle, and high school students were analyzed using Coh-Metrix. Text features, including text cohesion indexes, were described and analyzed regarding their changes depending on the material difficulty levels. For Text Analysis 2, texts in EIKEN reading sections were analyzed using Coh-Metrix. In addition to Text Analysis 1, textual features and their transitions according to difficulty levels of EIKEN are discussed.

Chapter 4 describes Study 2, which explored the effects of text cohesion (revised text in terms of cohesion) on Japanese EFL readers. Study 2 comprised six experimental studies (Experiments 1, 2-A, 2-B, 3-A, 3-B, and 4). Experiment 1 was conducted to examine whether Japanese EFL readers are sensitive to English text cohesion. A cohesion judgment task was used to measure readers' sensitivity to cohesion. Experiment 2 investigated the effects of the cohesion of expository texts on Japanese EFL readers' comprehension. Experiment 2 used different cohesive expository texts on social topics (high- and low-cohesion texts), and included different proficiency participants (relatively high-proficiency EFL readers for Experiment 2-A and relatively low-proficiency EFL

readers for Experiment 2-B), and compared their comprehension.

In the succeeding experiment, Experiment 3 testified the effects of the cohesion of science texts on Japanese EFL readers' comprehension. Experiment 3-A targeted relatively high-proficiency EFL readers, and Experiment 3-B, relatively low-proficiency EFL readers. Finally, Experiment 4 was conducted to investigate whether Japanese EFL readers changed their processing during reading according to text cohesion. Reader processing during reading different cohesive texts was assessed using a think-aloud task.

Given the results of two text analyses and six experimental studies, Chapter 5 discusses the findings of this study. I offer general discussion in terms of the following perspectives: (a) text difficulties and cohesion in reading, (b) cohesion manipulation and readers' understanding, (c) readers' factors affecting L2 reading.

Finally, Chapter 6 summarizes the findings. I highlight the importance of text cohesion in current English materials and its difficulty levels for Japanese EFL learners. Lastly, this thesis presents the limitations of the present study and suggestions and implications for future research on reading comprehension and material developers and educators of Japanese EFL learners.

Chapter 2

Review of Related Literature

2.1 Reading Comprehension and Processes

2.1.1 Theories and Models of Reading

Reading is one of the most important activities in acquiring knowledge or information. Thus, reading comprehension skill is essential for a person to be successful in educational and academic settings. Readers need to understand explicit and implicit meanings of texts and interact with the texts (writers). However, writers do not always convey the message explicitly enough. In such cases, readers are required not only to comprehend the explicitly stated information, but also to infer the implicit meaning between lines and construct a coherent mental representation of the text.

So far, research on text comprehension has explored how readers construct mental representations from texts. Researchers proposed several models to explain reading mechanisms (e.g., Albrecht & O'Brien, 1993; Gernsbacher, 1990; Graesser et al., 1994; Kintsch, 1998; Myers & O'Brien, 1998; van den Broek et al., 1996; van Dijk & Kintsch, 1983; Zwaan et al., 1995). One of the most influential models is the construction-integration model (CI model) proposed by van Dijk and Kintsch (Kintsch, 1988; van Dijk and Kintsch, 1983). They classified text comprehension level into three distinctive phases: (a) the *surface code* (i.e., memory for the surface linguistic structure of the text), (b) the *propositional textbase* (i.e., memory for the meaning that is explicitly stated in the text), and (c) the *situation model* (i.e., memory for events, states, and actions that occur in the text or ideas that are presented in the micro-world that the text describes).

The surface code is most likely to be forgotten. The propositional textbase and the situation model contribute to constructing the coherent representation of the text that is

regarded as successful comprehension. The propositional textbase and the situation model, which are more stable memory models compared to the surface code, consist of inferences generated from the readers' general knowledge. The situation model representation is the most stable comprehension model because it includes greater amounts of inferences. The situation model illustrates the deepest understanding of the text where inferences based on readers' knowledge are generated and integrated into the comprehension from the text. Moreover, the situation model involves deeper processing compared to the textbase since it incorporates information from readers' prior knowledge (e.g., Kintsch, 1988).

Readers' surface code or propositional textbase level of understanding is often measured by written recall tasks, such as answering literal comprehension questions of the text (e.g., true/false questions). On the other hand, the situation model level comprehension is measured by open-ended questions, inference questions, and so on. Thus, readers need to engage in various cognitive processes and accomplish the situation-model level comprehension. To construct a coherent representation, readers need to capture macro- and micro-structures of a text and comprehend then integrate information with their knowledge (Kintsch, 1994; Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983). In addition, as they encounter successive ideas and concepts when reading a text, this mental representation must be updated continuously (e.g., Kintsch & van Dijk, 1978, Rapp & Kendeou, 2007). Thus, successful reading requires the readers to dynamically apprehend and integrate each piece of information described or evoked in a text into an existing mental depiction or knowledge.

In other reading theories, readers' processing during reading has been discussed. McNamara and Magliano (2009) summarized similarities of the reading mechanism assumed in the major seven reading theories, namely the construction-integration model, the structure-building model, the resonance model, the event-indexing model, the causal

network model, the constructionist theory, and the landscape model (Albrecht & O'Brien, 1993; Gernsbacher, 1990; Graesser et al., 1994; Kintsch, 1998; Myers & O'Brien, 1998; van den Broek et al., 1996; Zwaan et al., 1995). In particular, the mechanisms are (a) connectionist architecture, (b) spreading activation, (c) automatic unconscious processing, (d) discourse focus, I convergence and constraint satisfaction, (f) mapping, (g) text-based inferencing, and (h) memory constraints. See Table 2.1 for the summary of the mechanisms by McNamara and Magliano. These are critical mechanisms to explain reading.

Of particular importance is that the input (i.e. textual information) that the reader receives while reading, affects the information that the reader activates and remembers. In addition, when there is a cohesion gap in the text, or when information mapping does not occur sufficiently, the reader processes inference generation. In this matter, reading is accomplished through a lot of complicated cognitive processes.

Table 2.1

Common and Shared Mechanisms Assumed in the Seven Reading Theories Summarized by McNamara and Magliano (2009)

Mechanism	Explanation
(a) connectionist architecture	- Comprehension involves the parallel activation of information in the environment (e.g., words in the text), the underlying meaning of that information, and prior knowledge. -Activation sources are often represented as layers in a network of nodes and links, with nodes representing words, propositions, or concepts, and links representing the relationships between them (e.g., predicates, verbs, causal connections).
(b) spreading activation	-The activation of concepts spreads activation to related concepts, resulting in a change in their activation. -Sorting of retrieval mechanism determines what information is initially activated or available. -Spreading activation is applied to the available concepts in memory, and this process changes the activation of concepts depending on their connectivity and initial strengths in the representation.
(c) automatic unconscious processing	- Some information is available automatically during reading, and there is some level of processing that is not consciously available to the reader.
(d) discourse focus	-There is an attentional focus by the reader and this focus changes across time and as the input changes. - The memorial strength of concepts and ideas is in part related to the amount of attentional focus they receive during encoding.
(e) convergence and constraint satisfaction	-The activation of any given concept or idea is based on the degree to which it receives activation from related concepts and ideas. -The mental representation is constrained by activated concepts and the relations between concepts in the input, as well as by information available from long-term memory.
(f) mapping	- Readers engage in unconscious processes to establish how the current linguistic input is related to the prior context. - This mapping is influenced by referential and situational cohesion. - Readers try to generate inferences when mapping fails.
(g) text-based inferencing	-Readers generate inferences (also called “bridging”) to establish connections between discourse constituents. -Bridging inferences can occur when mapping processes encounter referential or situational cohesion gaps. -Relationships between ideas in the text must be inferred when explicit cues such as argument overlap and connectives are absent. -These inferences may be considered part of the situation model to the extent that they reflect causal, motivational, temporal, and spatial relationships.
(h) memory constraints	-Readers’ working memory capacity is limited.

Note. This table was created based on the contents of Table 1 (p.304) in McNamara and Magliano (2009).

2.1.2 Text Comprehension and Processing: Constructing the Situation Model

As reviewed, text comprehension (e.g., Kintsch, 1988) is accomplished through numerous complicated cognitive processes. Additionally, it involves lower-level processes (e.g., word recognition, phonological processing, decoding words, analyzing the meanings of words and sentences), and higher-level processes (e.g., inference, association). The former are considered to support the readers' local understanding (e.g., the surface code) and depend on the bottom-up process of reading; however, the latter are regarded as necessary for the global understanding of a text (e.g., the situation model) and for grasping the complete text messages. Thus, the higher-level processes require the readers to perform top-down reading, such as inference generation and comprehension monitoring.

According to Grabe (2009), these cognitive processes are common in L1 and L2 reading. In general, these processes are carried out automatically in L1 reading, except under specific circumstances (e.g., memory tests). However, in L2 reading, readers often have limited reading skill, and they are not able to carry the above processes completely and automatically. In addition, it is suggested that L2 readers are not fluent in the lower-level processes; owing to this they have difficulty in engaging in higher-level processes such as monitoring comprehension or integrating separate parts of the text information (Grabe, 2009; Grabe & Stroller, 2019; Horiba, 2000). Thus, to read fluently, readers need to be good at lower-level processes.

Successful comprehension requires readers' engagement not only in the lower-level processes but also in the higher-level ones. As noted previously, inference generation to connect the text elements and thus grasp the text's global message contributes to the construction of a coherent representation (Graesser et al., 1994). Thus,

readers must process and interpret texts gradually with both bottom-up and top-down processes (Carrell & Eisterhold, 1983).

2.1.3 The Landscape Model

This section introduces the *landscape model* that explains coherent comprehension and spontaneous inferential behavior by readers (e.g., Linderholm et al., 2004; van den Broek et al., 1999). I introduce this model because it illustrates the dynamic processes assumed in the constructionist and memory-based theories.

The landscape model is characterized by the following two points: First, it incorporates the readers' specific purpose as an element that affects reading and is related to the *standards of coherence* (as will be explained later) that determine the information needed for reading. Second, this model includes the readers' use of background knowledge to achieve a coherent comprehension. Previous empirical studies have evaluated the landscape model as having a certain validity (van den Broek et al., 1996; van den Broek et al., 1999).

The landscape model assumes that reading involves fluctuating concepts activated in a cycle (e.g., Gerrig & McKoon, 1998; Linderholm et al., 2004). The model assumes four potential sources: (a) *current text* (the part of a text that provides the readers with new input), (b) *information transferred from the prior reading cycle* (readily available information previously read or activated by the readers), (c) *reinstated information* (relevant information activated in a previous reading cycle), and (d) *readers' background knowledge* (information that the readers possessed prior to reading the texts).

In this model, new information is processed and activated in the following sequence: First, the current text provides an original input to the readers that they process or activate. Second, the information activated in the prior reading becomes readily

available and is transferred. Third, the readers reinstate the information activated in a previous reading cycle; specifically, the information that is not readily available is activated to understand the text. The reinstated information may include the actual text information and other data. Lastly, the readers may activate background knowledge to access their semantic memory to comprehend the text.

Moreover, this model involves readers' standards of coherence. The standards of coherence reflect the degree of comprehension that a reader attempts to attain during reading (Linderholm et al., 2004; van den Broek et al., 1995). In other words, they comprise the types and strengths of coherence that the readers try to achieve while reading texts (Linderholm et al., 2004; van den Broek et al., 1995; van den Broek et al., 2005). They are determined by several factors, such as the reader's purpose (e.g., Linderholm & van den Broek, 2002; van den Broek et al., 2001; for study, for entertainment), background knowledge, text difficulty, distraction during reading, fatigue, and comprehension strategies (e.g., Graesser et al., 1994; van den Broek et al., 1995). For example, a stronger degree of coherence is required when reading for learning than for entertainment. Thus, these standards are considered as inter- and intra-individual factors in reading (Nahatame, 2017) and they dictate the sources that should be activated during reading (Linderholm et al., 2004).

Based on the standards of coherence, the readers attempt to grasp the text's meaning. When they have high standards, they attempt to understand the text both in detail and globally. However, when they have low standards, they do not force themselves to comprehend every detail; they only attempt to grasp the outlines or the necessary information (Linderholm et al., 2004).

According to previous studies, readers can monitor various coherence types, such as referential, causal, intentional, protagonist, spatial, and temporal coherence (cf. the

event indexing model of reading; Sanders et al., 1992, Zwaan et al., 1995). In particular, referential (i.e., text cohesion) and causal coherence have been regarded as its critical types in reading (e.g., van den Broek et al, 2002). It has been demonstrated that both native and second language learners possess the standards of coherence regarding the meaning and causal coherence (Nahatame, 2017).

Further, the landscape model explains two different mechanisms are assumed in terms of information use under reading: *cohort activation* and *coherence-based retrieval*. The former consists of the extending of activation from the currently stimulated to the related information in the mental representation or background knowledge. It is fast and passive and a memory-based mechanism (cf. the resonance model; Myers & O'Brien, 1998). The second type of mechanism, coherence-based retrieval, is a strategic and deliberate retrieval to construct coherence for understanding the current text, such as reviewing a previously read section to access prior information or representation (Goldman & Saul, 1990). Furthermore, it entails the readers' knowledge-based retrieval, which is slow and laborious (e.g., Linderholm et al., 2004).

Therefore, the landscape model comprises two types of activation (i.e., bottom-up and top-down processes) of information (i.e., cohort activation and coherence-based retrieval, respectively). The type of retrieval employed depends on the context. If the activated information is sufficient to correspond to the readers' standards of coherence, the connection between the concepts occurs easily and speedily by passive, autonomous processes (i.e., cohort activation). However, when it is insufficient to satisfy the readers' standards, then more demanding, strategic processes (i.e., coherence-based retrieval) are initiated to establish an association between the pieces of information to achieve the desired level of coherence. Thus, the standards of coherence play a critical role in

determining what processes the readers need to engage in during reading. In the following two sections, reader and text factors affecting reading will be reviewed.

2.2 Factors of Readers and Texts in Reading

2.2.1 Reader Factor: Language Knowledge

As reviewed, reading involves many cognitive processes. One of critical factors that affect reading quality is readers' language knowledge. In both L1 and L2 reading, lexical knowledge is critical in interpreting text meaning, and readers need to recognize words visually. This means using the visual and phonetic systems while decoding the constituents of the lexicon through grapheme-phoneme matching (Chen et al., 2015). Thus, to comprehend a text, readers need to have phonological, morphological, and semantic information of vocabulary, and parse a sentence with grammatical knowledge. Previous research suggested that reading and lexical knowledge are mutually related (e.g., Freebody & Anderson, 1983).

2.2.2 Reader Factor: Proficiency and Inference Skills

When the readers comprehend scripts, they must perform many cognitive activities, decode text information, draw inferences, integrate the text and their own knowledge, and establish and maintain a coherent mental representation of the text. Previous studies have investigated the inferencing process in reading and demonstrated the importance of higher-level reading processes in accomplishing deep comprehension in reading (Graesser et al., 1994; van Dijk & Kintsch, 1983).

According to van den Broek (1994), inference is defined as "information that is activated during reading yet not explicitly stated in the text" (p. 556). Although prior studies have presented its many types (e.g., Graesser et al., 1994; van Dijk & Kintsch,

1983), in this research, we focus on the *bridging inference*. Bridging inference is needed to connect the focal statement with the previous text and establish comprehension coherence. It is regarded as essential to construct coherent mental representations, and by generating it, the readers can fill in the gaps between sentences (e.g., Myers et al., 1987). Prior research has shown that some inferences produced during reading improve reading comprehension (e.g., bridging inference); moreover, good readers are more skilled at generating them (e.g., McMaster et al., 2012; McNamara & Kintsch, 1996).

However, in some cases, it is challenging for the readers to study texts and construct coherent mental representations. These difficulties can arise from their inability to infer (e.g., Perfetti, 1985). Poorly-skilled readers make fewer inferences than the skilled ones because inferencing is a high-load cognitive activity. The former tend to engage in lower-level processes (Horiba, 1996, 2000), whereas the latter, who are able to capture the meaning without that burden, do not have to.

Inferencing has been perceived as a demanding process (Cain & Oakhill, 1999; Oakhill, 1984) and insufficient inferencing skills may cause difficulty in reading (Kintsch & van Dijk, 1987; Yuill & Oakhill, 1991). Previous L2 studies have examined the L2 readers' online reading process using the online method (i.e., think-aloud task) that measures the allocation of cognitive resources by collecting their verbal reports during reading (Horiba, 1996, 2000; Kimura, 2015; Shimizu, 2009, 2015). It has been reported that these readers have to devote their cognitive resources to the lower-level processes (i.e., word recognition, syntactic analysis, meaning construction); moreover, they were unable to generate inferences, contrary to the L1 readers (e.g., Horiba, 1996). Additionally, L2 reading proficiency has been found to influence the number of inferences (Shimizu, 2009).

In summary, generating inferences is critical in the construct situation model.

However, due to the high cognitive load, L1 and L2 readers often have problems in generating inferences (e.g., Horiba, 1996, 2000; Kintsch & van Dijk, 1978).

2.2.3 Text Factor: Cohesion and Coherence

As mentioned, there are several important reader factors (e.g., the readers' standards of coherence and skills of inference generation) affecting the success of reading. In addition to them, the text factors also affect reading and comprehension. This is because a reader's mental representation is constructed based on the text input (e.g., Linderholm et al., 2004; McNamara & Magliano, 2009); if the script provides insufficient information, the reader will fail to achieve a coherent comprehension. In this section, we focus on two important concepts in text comprehension: *text cohesion* and *coherence*.

Since the focus of linguistics and reading research has been on text as a written language, text cohesion and coherence have been important notions that have attracted the attention of scholars (e.g., Carrell, 1987; Halliday & Hasan, 1976; McCrudden & McNamara, 2017; McNamara & Kintsch, 1996).

According to Halliday and Hasan (1976), leading researchers in the field, cohesion is defined as "relations of meaning that exist within the text, and that define it as a text" (p.4). It is what makes a text meaningful (cf. texture) and occurs when the interpretation of one of its elements depends on the other. According to this definition, there are two kinds of cohesion: grammatical, such as indication, substitution, and omission, and lexical, such as repetition. The conjunctive expressions exist as an intermediate between the two

Recent L1 and L2 studies have used the term text cohesion (cohesiveness) to refer to how connected and explicit a text's linguistic elements are (e.g., Graesser & McNamara, 2011; Hosoda, 2016; McNamara et al., 2014). Text cohesion refers to the degree to which

the concepts, ideas, and relations within a text are explicit, and text coherence entails the effect of text cohesion on the readers' comprehension (Graesser et al., 2003). Although text cohesion and coherence are related concepts, they are regarded as different aspects of a text (Li & D'Angelo, 2015). According to Warren (2012), the latter is defined as the manner in which the text elements, such as sentences or phrases, are associated in terms of their meaning. Although text cohesion contributes to textual coherence, it is not always necessary for it. That is, highly cohesive texts containing connective expressions can help the reader integrate sentences easily; however, readers can infer the relation of sentences even without such cues, based on their text-based inferences (McNamara & Magliano, 2009) and coherence-based retrieval during reading (e.g., Linderholm et al., 2004). Thus, it would be possible for a script to be highly cohesive in terms of lexical and discourse level, however, insufficiently coherent, and vice versa (Traxler, 2011).

Recent studies (e.g., Graesser et al., 2004; McNamara et al., 2014) have described coherence as a characteristic of a text's mental representation. Specifically, while cohesion consists of explicit wordings (e.g., connectives, argument overlapping) as described above, coherence is defined as the type of connectedness that the readers infer among the propositions in the discourse representations. Briefly, it can be derived from both the text information and the reader's prior knowledge and is generally understood in terms of meaning. It is important to note that text cohesion and coherence have been confounded occasionally, and previous studies have used one term to describe the other (e.g., McNamara et al., 1996).

2.3 Effects of Text Cohesion on Reading

2.3.1 Definition of Text Cohesion

Previous empirical studies have reported that the less skilled readers are unable to infer a text's implicit message when the target information is inexplicitly presented (Linderholm et al., 2000; McNamara & Kintsch, 1996). Disconnected information without sufficient explanation to facilitate a text's coherent understanding impedes the readers' comprehension by compelling them to construct superficial and fragmented representations (O'Reilly & McNamara, 2007).

With the aim of rendering these complicated and burdensome reading processes easier for the readers, prior research has investigated the text factors (e.g., Linderholm et al., 2000; McNamara, 2001; McNamara et al., 1996; McNamara et al., 2011; Vidal-Abarca et al., 2000). *Text cohesion* is an important text feature that affects the readers' comprehension. Cohesion can be changed using linguistic cues (e.g., connectives such as *however* or *therefore*) to join adjacent sentences (Ozuru et al., 2009) by using the same words or phrases to make texts more explicit (e.g., *argument overlapping*) or by adding headers and topic sentences (McNamara et al., 1996).

Text cohesion is necessary for the construction of coherent mental representations (O'Reilly & McNamara, 2007). Furthermore, it helps the readers maintain comprehensible representations of the text by reducing the need for reader inference and/or the text's integration with the readers' knowledge (Ozuru et al., 2009). When the readers study a high-cohesion text, they can gain a sufficient amount of information for comprehension from the text itself. However, when reading a low-cohesion text, they have to infer necessary ideas to cover the gaps in it (Hosoda, 2016; Kintsch, 1994; McNamara et al., 1996; McNamara & Kintsch., 1996; Ozuru et al., 2010). Thus, text

cohesion is considered to be a significant aspect of text difficulty (readability index; McNamara et al., 2012).

2.3.2 Effects of Text Cohesion in L1 Reading

To resolve the difficulty of reading, previous studies in the L1 reading field have attempted to make texts easier to comprehend by revising them to be more explicit or informative (e.g., Gilabert et al., 2005; Hall et al., 2014; Hall et al., 2015; Kleijn et al., 2019; Linderholm et al., 2000; McNamara, 2001; McNamara et al., 1996; McNamara et al., 2011; Ozuru et al., 2009; Vidal-Abarca et al., 2000). In these experimental studies, text cohesion was manipulated (e.g., adding/deleting repetition of words, phrases, and connectives), and the readers' comprehension was measured using post-reading tasks (e.g., written recall, multiple-choice problems, open-ended questions). Most of them revealed that high-cohesion texts are helpful for the readers, especially those who find reading challenging (Beck et al., 1991; Linderholm et al., 2000; Vidal-Abarca et al., 2000).

For example, empirical research by Loxterman et al. (1994) reported that as compared to the low-cohesion social-study texts, the high-cohesion ones improved the L1 readers' text memory (recall amount). These results indicated that the low-proficiency readers who draw inferences poorly can take advantage of the explicit text signals from the high-cohesion texts to fill in the gaps in them and smoothly construct its coherent representations.

In addition, Best et al. (2006) confirmed the positive effect of high-cohesion narrative scripts on the L1 English fourth-grade elementary students' global comprehension of a text, measured using multiple-choice questions. Similarly, Gilabert et al. (2005), who examined the impact of historical texts' cohesion on junior high school students and undergraduates, reported that high-cohesion texts fostered the readers'

memory. The participants' recall and inference generation were assessed using open-ended and multiple-choice questions.

Hall et al. (2014) assessed the cohesion effect on the L1 secondary school students' comprehension of science texts measured using multiple-choice questions. They reported that high-cohesion texts (with repetition of words and phrases) supported the students' achievement of improved text comprehension. Hence, previous studies involving young native readers have revealed the positive outcome of revising cohesion of texts in history, social studies, and narrative genres. Specifically, they reported that raising text cohesion improved the young L1 readers' reading.

Regarding the adult L1 readers, Ozuru et al. (2010) investigated the influence of text cohesion on the L1 undergraduates' comprehension of science texts. They created high-cohesion texts in the following measures: (a) adding connectives (e.g., because), (b) replacing pronouns with specific noun phrases, (c) adding nouns to increase argument overlap between adjacent sentences, and (d) adding relational pronouns to make explicit the relations between phrases in a sentence. They measured readers' understanding of low- and high-cohesion texts through self-explanations (i.e., verbal descriptions produced to demonstrate and facilitate coherent understanding of the text during reading) and open-ended questions. They indicated that as compared to the low-cohesion texts, the high-cohesion ones prompted higher-quality-bridging inferences by the readers in self-explanations. Moreover, some studies have reported different effects on text cohesion. Ozuru et al. (2009) examined the effect of revising cohesion on the L1 undergraduates' comprehension of science texts using open-ended inferential questions and reported that the high-cohesion texts improved their textbase-level comprehension. In this way, the effectiveness of increasing the text cohesion of science texts has been demonstrated in L1 adult reading.

As a study with a slightly different cohesion modification, Linderholm et al. (2000) investigated the impact of revising the causal structure of difficult and easy texts for 39 undergraduate native speakers. In their experiment, the texts were revised to be more cohesive in three ways: (a) arranging the text events in temporal order, (b) making implicit text goals explicit, and (c) repairing the coherence breaks caused by inadequate explanations or distant causal relations. The results showed that regardless of the readers' proficiency, the revised high-cohesion texts enhanced their understanding in a recall task and on comprehension questions, as compared to low-cohesion texts that were difficult to read. These findings indicated a possibility of the existence of an interactional benefit of revising texts for the low-proficiency readers. Specifically, the high-cohesion texts were effective for these readers.

As noted above, text cohesion has been examined in many empirical studies (e.g., Gilabert et al., 2005; Hall et al., 2014; Hall et al., 2015; Linderholm et al., 2000; Ozuru et al., 2009; Vidal-Abarca et al., 2000); moreover, the results have indicated that high-cohesion texts, which eliminate the readers' burden of predicting the relationships that remain unexplained by the gaps in them, had positive effects on their comprehension. As a tendency, various types of texts has been used for young readers, but science texts have been used for adult readers to prove the effects of cohesion manipulation on L1 reading. This is reasonable, considering that cohesion plays an important role in reading, where readers comprehend unfamiliar contents such as science expository text (McCrudden & McNamara, 2017; Schmitz et al., 2017).

In more recent studies, the effects of revising cohesion have been confirmed. Désiron et al. (2020) extended the previous findings of cohesion effects to long multimedia-document reading. They examined the effects of text cohesion and explicit signaling on a five-page-long multimedia document about river sailing, by measuring

readers' comprehension using open-ended questions and processing by eye tracking. They created highly cohesive texts based on Ozuru et al. (2009) and signaled multimedia documents by adding captions and arrows in the picture, and using the same color-coding in texts. The results suggested that young students (around 17 years old) tended to understand the high-cohesion texts better than the low-cohesion ones (Experiment 1). While college-level students did not have a dominant effect on comprehension, their attention allocation during reading was different between text conditions (Experiment 2). Specifically, they concentrated more on the images contained in the text when they read the enhanced signaling text. It was suggested that the high-cohesion texts may be effective for the less proficient readers' comprehension and that the degree of textual clarity may affect readers' attention allocation, even for adult readers in multimedia documents.

Additionally, Schmitz et al. (2017) investigated how textual features (local/global cohesion) impact junior high school students' understanding. They also investigated whether the readers' comprehension of text differs when they expect a text to be expository or literary (genre expectation), considering their proficiency and background knowledge. They manipulated text cohesion by adding pronouns (local cohesion), paragraph titles, proposition summarizing sentences, and explicit relation expressions (global cohesion) based on McNamara and Kintsch (1996). The results suggested that the global-cohesion text facilitated readers' comprehension, measured by multiple-choice and open-ended questions when the reader assumed the text to be expository. This study suggested that the effect of cohesion manipulation can vary depending on the genre of the text assumed by the reader.

2.3.3 Reverse Cohesion Effects in L1 Reading

However, it has been argued that reading high-cohesion texts is seldom facilitative for all readers. McNamara et al.'s studies have pointed out that such texts may not benefit the high-knowledge readers (McNamara, 2001; McNamara et al., 1996; McNamara et al., 2011; O'Reilly & McNamara, 2007; Smith et al., 2021). This is called the *reverse cohesion effect*, a counterintuitive finding that the high-knowledge readers gain an enhanced understanding from the low-cohesion texts. McNamara et al. (1996) examined the role of text cohesion in biology texts on the comprehension of 36 L1 junior high school students and measured it using a written recall task. They found that the readers with limited domain knowledge of the text benefitted from a high-cohesion text; nevertheless, the high-knowledge readers found the low-cohesion text to be advantageous.

Similarly, a high-cohesion text was significantly efficacious for the low-knowledge college readers, however, not for the high-knowledge ones (McNamara, 2001; McNamara & Kintsch, 1996). This is because the less connected the information, the more inference generated by the reader, and the more the facilitation of the text memory (e.g., Dahl et al., 2020; Nahatame, 2013; Radvansky et al., 2014).

McNamara et al. (1996) offered a possible explanation for these reverse cohesion effects: high-knowledge readers do not tend to access and use their own knowledge to infer when they are given a high-cohesion text. In such a text, the readers' inference generation might be inhibited because it is clearly described and therefore does not require their reasoning. Thus, when the readers have sufficient knowledge needed to construct coherent mental representations, a highly cohesive text is unnecessary. To be more precise, high-cohesion texts tend to make the readers more passive, such that they do not tend to infer, associate, or monitor the text (McKoon & Ratcliff, 1992). Therefore, low-cohesion texts render some readers active to the point that they generate inferences

to connect the gaps within the text; moreover, their knowledge and text contents are integrated for an improved comprehension (O'Reilly & McNamara, 2007).

Similarly, Ozuru et al. (2010) reported that low-cohesion texts induced the L1 undergraduates to draw inferences to fill conceptual gaps in challenging texts; they achieved enhanced comprehension measured using open-ended questions when reading a low-cohesion text as compared to a high-cohesion one. Furthermore, McNamara et al. (2011) investigated the effects of cohesion on the L1 elementary students' comprehension of narrative and science expository texts; they reported the reverse cohesion effect in which the high-knowledge readers understood the low-cohesion text better than the high-cohesion one. In addition, Radvansky et al. (2014) also indicated that texts with causal breaks (i.e., low-cohesion texts) prompted better memory of the script in the readers.

Nevertheless, it has been suggested that the reverse cohesion effect may vary depending on the readers' proficiency (e.g., O'Reilly & McNamara, 2007). O'Reilly and McNamara (2007) examined the effect of text cohesion on the L1 college students' reading and found that the benefit of low-cohesion text was restricted exclusively to the low-proficiency readers with high knowledge. However, in addition to this effect, they also reported that the readers having high levels of proficiency and knowledge were able to comprehend the high-cohesion text better than the low-cohesion one. Further, Hall et al. (2015) demonstrated similar results. They conducted an empirical study investigating the impact of cohesion on the L1 secondary school children's text comprehension assessed using multiple-choice questions. They revealed that the students comprehended high-cohesion text more effectively; however, the high-proficiency readers performed satisfactorily regarding both high- and low-cohesion texts.

Kleijn et al. (2019) focused on the types of conjunctions that manipulate cohesion and examined how the presence or absence of coherence marking affects the reader's

understanding of the text. Overall, 94 junior high school native speakers were provided a text in which causal, contrastive, additional, or temporal conjunctions were either added (i.e., high cohesion) or deleted (i.e., low cohesion). Comprehension was measured using the cloze tests. The results showed that high cohesion was more effective than low cohesion in understanding texts with manipulated causal and contrastive connectives. In the case of texts manipulating additional connectives, low cohesion may be more helpful than high cohesion for local comprehension. Furthermore, it was suggested that these effects of cohesion revision may affect the understanding of the passages other than those in which the conjunctions were manipulated, especially in more challenging texts. This study suggested that connectives in texts can affect young readers' comprehension that may be related to the text difficulty.

In summary, high-cohesion texts, where ideas and information are described explicitly, have demonstrated positive effects on the readers' comprehension in many studies (e.g., Best et al., 2006; Gilabert et al., 2005; Hall et al., 2014; Linderholm et al., 2000; Loxterman et al., 1994; Ozuru et al., 2009; Ozuru et al., 2010). Nevertheless, others have reported a reverse cohesion effect whereby low-cohesion texts facilitate the comprehension of those readers with high proficiency or knowledge (McNamara, 2001; McNamara et al., 1996; McNamara et al., 2011; O'Reilly & McNamara, 2007). Although inconsistent results have been reported (Hall et al., 2015; O'Reilly & McNamara, 2007), prior studies have reached the following consensus: text cohesion (high or low) affects the L1 readers' comprehension.

Thus, the research on native language reading has verified some effect of cohesion on reading comprehension; nonetheless, it varies depending on the readers' characteristics and level of knowledge. Additionally, it has been suggested that the impact on comprehension may differ depending on the type and nature of the inter-sentential

connections being modified (Kleijn et al., 2019; Murray, 1997); the effect may also vary according to the text genre (Kamalski et al., 2008; Schmitz et al., 2017).

2.3.4 Effects of Text Cohesion in L2 Reading

In addition to the studies on L1, those on L2 have also investigated the effect of text cohesion on reading (e.g., Carrell, 1987; Farrokh & Gavabari, 2019; Horiba, 1996, Hosoda, 2016). However, the latter is exceedingly limited as compared to the former. A pioneering study by Horiba (1996) focused on the causal connections in a script (Trabasso & van den Broek, 1985) and investigated the influence of different cohesive texts on the readers' text comprehension and processing. Specifically, she targeted the Japanese (n = 20) and English native speakers (n = 56 with different Japanese proficiency levels). By reducing the number of causal links between the events in the transcript, she created experimental texts written in Japanese (in her study, she called them high- and low-coherence texts). She measured the participants' text comprehension and processing of the diverse coherence texts using a recall task and verbal reports, respectively. The results showed that the native speakers of Japanese engaged in higher-level processes (i.e., general inference, use of general knowledge association) during reading, and they demonstrated discrepancies regarding processes in the different coherence texts: the participants created more elaborations for the low-coherence texts than for the high-coherence ones. Contrariwise, those who read the experimental texts as their L2 allocated their attention to the lower-level processes and did not process differently between the high- and low-coherence texts. However, in this study, she focused only on the manipulation of causality and did not examine the effects of text cohesion, including text redundancy and explicitness.

In a subsequent study, Horiba (2017) explored the influences of the L2 readers'

linguistic knowledge and text cohesion on expository text understanding. She especially targeted the English- and Chinese-speaking learners with Japanese as a second language as well as the Japanese native speakers. She formulated two cohesion texts by eliminating explicit markers in them to lower the text cohesion (in low-cohesion text) and employed a recall task to measure the readers' comprehension. The results showed no significant differences in the recall amount between the high- and low-cohesion texts. She insisted that the L2 readers did not benefit from the explicit expressions when reading the high-cohesion texts. However, she analyzed the descriptive data and discussed that the L2 readers recalled more information when reading the low-cohesion texts, as compared to the high-cohesion ones; nevertheless, the contrary was true for the native speakers. It should be noted, however, that these findings are from an interaction that demonstrated a statistically significant trend (p = .056) but not a statistically clear difference. This is an interesting study that focuses on text cohesion and reader characteristics. However, since it was Japanese reading comprehension, it may be difficult to apply it directly to English reading comprehension.

In addition, as a study on the Japanese EFL readers, which is the focus of the present study, Hosoda (2016) conducted an experimental study on Japanese EFL undergraduates and employed three tests (i.e., written recall, causal questions, and a problem-solving assessment) to reveal the impact of the interaction effects of text cohesion and readers' proficiency. He used high- and low-cohesion expository texts about physical phenomena based on the cohesion manipulation of Ozuru et al. (2010) and used several tasks to measure the participants' comprehension in different aspects (textbase and situation model comprehension). The results showed that text cohesion prompted the readers' text memory regardless of their L2 proficiency. Moreover, a reverse cohesion effect was observed in the understanding of the relations in the text: the high- and low-

proficiency readers performed better on the tests when reading the low- and high-cohesion texts, respectively. This research concluded that text cohesion affects the readers' textbase-level comprehension more than the situation-model-level comprehension.

However, in contrast to the reverse cohesion effect confirmed by Hosoda (2016), a recent study by Farrokh and Gavabari (2019), who investigated manipulating text cohesion in adult Iranian EFL readers, revealed that compared to the low-cohesion texts, the high-cohesion ones improved comprehension.

In summary, although previous L2 studies have not reached an agreement (i.e., on the normal or the reverse cohesion effect), they have reported that text cohesion can affect readers' processing and comprehension, similar to the L1 research. However, there are limited studies investigating the influence of text cohesion, as compared to the extensive body of the L1 research. Furthermore, although there are studies (e.g., Nahatame, 2017) investigating readers' strength of standards of coherence (coherence judgment) in L2 reading, limited information is available regarding how L2 or EFL readers monitor and react to text cohesion during reading. Hence, it would be desirable to reveal whether Japanese EFL readers examine text cohesion during reading.

2.3.5 Text Analysis Regarding Cohesion in the L2/EFL Research

Based on the role and importance of text cohesion in reading comprehension reviewed thus far, recent research on reading and teaching materials has analyzed texts from the perspective of text cohesion.

Previously, the supposed *shallow-based readability formulas* (e.g., the Flesch-Kincaid Grade Level formulas, Klare, 1974–1975) have been employed as an indicator of text difficulty. This is based on word frequency and the length of the sentence; we assume that text comprehension is relevant to the possibility that a reader knows the

words and interprets the sentences in the text. If a text contains longer words that are often less frequent in English usage, the text is considered to be difficult.

Although these metrics of text difficulty can be well-established and reasonable, considering that word and sentence length are good predictors of reading ease (Just & Carpenter, 1987), this single measure cannot reflect multiple facets, such as word meaning, concreteness, and discourse relations, of text comprehension (McNamara et al., 2014).

However, as an alternative readability formula to determine text difficulty, text cohesion has attracted researchers. As reviewed previously, researchers in linguistics, L1 and L2/EFL reading, and discourse processing have insisted on the significance of text cohesion in reading (e.g., Halliday & Hasan, 1976; Graesser et al., 2003; van Dijk & Kintsch, 1983). This is because text cohesion is defined as how textual elements are related and the extent to which sentences have overlapped expressions. It can explain readers' burden to infer relations within the text (or text difficulty). As cohesive devices (i.e., discourse markers, connectives, coherence markers, or signals) are essential factors for constructing meaningful representations of text comprehension, material developers must understand and make use of them, to develop good educational texts that facilitate language learners' reading skills. With the reasons reviewed above, text cohesion is considered to be a significant aspect of text difficulty (readability index; McNamara et al., 2012). Hence, McNamara and colleagues developed Coh-Metrix (Graesser et al., 2004; McNamara et al., 2014) to scale text difficulties.

Coh-Metrix is designed to provide multiple features of the text, including words and syntax, textbase and situation model, discourse genre and rhetorical structure, the pragmatic communication level, and so on (e.g., Graesser & McNamara, 2011). Coh-Metrix specifically focuses on text cohesion ("Coh" stands for cohesion), to establish an

objective means to scale text cohesion. They assumed (a) text cohesion is in the text (represented by text elements) and (b) text cohesion can be computationally measurable. Since text cohesion emerges from the absence/presence of cohesive cues (e.g., connectives) in the text, Coh-Metrix determines the cohesiveness of the text in terms of the frequencies of the cohesive cues in the text. In particular, they established the following two indexes for cohesion in their easability assessor.

> *Referential Cohesion. A text with high referential cohesion contains words and ideas that overlap across sentences and the entire text, forming explicit threads that connect the text for the reader. Low-cohesion text is typically more difficult to process because there are fewer connections that tie the ideas together for the reader.*
>
> *Deep Cohesion. This dimension reflects the degree to which the text contains causal and intentional connectives when there are causal and logical relationships within the text. These connectives help the reader form* [sic] *a deeper and more coherent understanding of the causal events, processes, and actions in the text. When a text contains many relationships but does not contain those connectives, the reader must infer the relationships between the ideas in the text. If the text is high in deep cohesion, then those relationships and global cohesion are more explicit.* (p.85, McNamara et al., 2014)

Coh-Metrix can be applied to various types of text (e.g., naturalistic texts, manipulated texts for experiments, conversations), and it can capture the textual features (McNamara et al., 2014). McNamara et al. suggest that text assignments to language

students should be within their optimal zone of comprehension difficulty. That is, text that is neither too easy nor too challenging for the students—considering their reading skills and proficiencies—is desirable (McNamara et al., 2014).

As this tool provides various perspectives from which to evaluate the lexical properties, syntactic features, and text cohesion of English texts, Coh-Metrix has been widely used in the analyses of texts written for native English speakers (e.g., McNamara et al., 2014). More recently, it has also been applied to analyze teaching materials for English learners (e.g., Nahatame, 2021; Nahatame & Kimura, 2019). Here, I review related L2/EFL research that used Coh-Metrix text analysis.

For example, Crossley and colleagues have examined the reading materials for the English as a second language (ESL) learners in terms of lexical, sentence, and text cohesion features by using Coh-Metrix (e.g., Crossley et al., 2012; Crossley et al., 2007). The results revealed that the simplified texts contain a greater coreferential cohesion and common connectives than the difficult ones. This implies that as the text difficulty increases, the cohesion based on the lexical and semantic overlap decreases, and reading comprehension becomes more difficult, as indicated in previous studies. However, it has also been confirmed that indicators such as conjunctions do not necessarily yield consistent results with respect to text difficulty.

In addition, Crossley et al. (2014) investigated the effects of text cohesion and lexical features on the source text integration in a speaking test. They assessed the text features of the source scripts in the speaking section of the Test of English as a Foreign Language – Internet-based test (TOEFL-iBT) and speaking responses of 480 examinees. They found that the textual features of the source scripts were strong predictors of the test takers' verbal responses (integrated words from the source texts) in the speaking test. Although these results are for the listening contexts, it is an important finding that features

such as the source texts' cohesion can affect the learners' performance as an outcome of understanding the input.

Similarly, Plakans and Zeynep (2016) examined text cohesion in reading passages in 27 college-level textbooks for ESL learners using a Coh-Metrix text analyzer. Furthermore, to explore the differences in the textual features in various grade-level texts, they analyzed the ESL textbooks from different levels (beginning, intermediate, and advanced) using multiple indexes calculated by Coh-Metrix. They found significant discrepancies in the text cohesion in the textbook passages of varying levels. More specifically, reference cohesion, defined as the proportion of the content words that overlap between the adjacent sentences, increased when comparing the beginning and intermediate, beginning, and advanced level textbooks. Correspondingly, the number of causal connectives contained in the textbooks increased when the text level was higher (beginning and intermediate, beginning and advanced pairs). However, such differences in the text cohesion in different textbooks were unverified in the levels adjacent to them: intermediate and advanced. Although the findings did not show that text cohesion increased as the grade level of the text became more difficult for all grades, this study is valuable in terms of researching the cohesion aspect of the ESL materials in college-level texts.

Moreover, Han and Shin (2016) aimed to gain empirical evidence for the utility of automatic text analyzers and investigate the role of textual features in an intensive English reading assessment program. They utilized textual characteristics such as lexical, semantic, and syntactic features calculated by some automatic text analysis tools (e.g., Coh-Metrix, Lexile); moreover, they explored whether these features account for the reading item difficulty in intensive English programs. The results revealed that some lexical characteristics (i.e., word length, total word count, lexical–semantic analysis, and

connectives) accounted for 45% of the variance in the reading item difficulty. Furthermore, they recommend that future studies employ text analyzers to examine various types of text in educational contexts, to obtain a more comprehensive view of the relationship between reading difficulty (or reader ability) and text characteristics.

In addition, there is research that has considered genre in the analysis of teaching materials. Roman et al. (2016) explored whether the logical connectives (e.g., therefore) used in the L2 textbooks differed between 12 social studies and science textbooks each. The results demonstrated that there was a higher rate of logical connectives in the latter than in the former. Moreover, they indicated that the types of textbooks interacted with logical connectives: the higher the grade level of the science textbooks, the greater the logical connectives used; however, such a result was unidentified in the social studies textbooks. This research suggested that the cohesive devices in texts may vary depending on their genre. However, it should be noted that the number of texts analyzed in Roman's study was limited. Furthermore, they suggested that "more studies should be concluded in ecologically valid contexts (e.g., schools)" (p. 14); that is, text analyses in terms of text cohesion should be conducted on a greater number of school textbooks.

To summarize, the ESL and EFL studies have analyzed English texts in the readers' tests and textbooks vis-à-vis their cohesiveness and grade level; they found a relationship between text difficulty and cohesiveness indicators. Nevertheless, the findings vary; moreover, it seems that some text cohesiveness indices show a higher cohesiveness for easier texts, while the others, such as conjunctions, indicate it for the more difficult ones. The question regarding the results of the studies with Japanese EFL learners remains unanswered.

Unfortunately, thus far, most studies that have been conducted on foreign language (English) textbooks, and materials have focused on the text content, such as the

percentage of the tasks and skills included (e.g., Tabata & Masuda, 2018; Torikai et al., 2017). Moreover, limited text cohesion analyses have been performed compared to the body of the L1/ ESL research. However, several studies have assessed the text cohesion of teaching materials for Japanese EFL readers (e.g., Nahatame, 2017; Nahatame, 2018; Nahatame & Kimura, 2019).

As for Coh-Metrix, an objective English analysis program that considers the text cohesion, Mori (2018) analyzed the English text of stories in the Japanese elementary school materials “Hi, friends!” and “We can!”. Moreover, Mizumoto (2013) examined the high school English textbooks in Japan and confirmed that the difficulty level of English sentences increases with the progression of the textbook. He argued that the use of Coh-Metrix makes it possible to conduct a more precise analysis than the conventional difficulty estimation. However, they only analyzed a limited volume of texts and did not present comprehensive findings for text cohesion in EFL materials.

In a more recent study, Nahatame (2021) assessed the text cohesion of the English reading materials (with different difficulty levels) for young children using the Tool for the Automatic Analysis of Cohesion (Crossley et al., 2016) to clarify the relationship with difficulty level. The results of the analysis of 129 Oxford Reading Tree books revealed that about 37% of the variance in readability was explained by four indicators: verb redundancy between sentences, semantic relatedness between sentences, conjunctive phrases related to time and order, and repetition of content words within sentences. Overall, the difficulty level tended to be lower for materials in which it was easy to construct content connections through overlapping vocabulary and their meanings and for those with fewer connective phrases. Specifically, this research suggested that the factors related to text cohesion, such as word redundancy and conjunctions, may also be associated with the level of text difficulty for young EFL learners.

However, these studies have been limited to a certain type of school and have not analyzed large text databases (e.g., textbooks from elementary to high schools) and clarified how textual characteristics, including text cohesion, have changed according to the different grade levels.

2.3.6 Text Cohesion and Instruction

Thus far, we have considered the effects of text cohesion on reading comprehension and studies that have examined text cohesion and text difficulty. Based on these previous studies, several others have emerged that have utilized text cohesion in the L2 and EFL instruction as well as text revision (e.g., Aidinlou et al., 2012; Safaie, 2020; Zhang, 2018). For example, Aidinlou et al. (2012) investigated the influences of a textual cohesive reference instruction on 60 Iranian EFL readers' comprehension; they implemented the instruction on half of them (the other half did not receive specific training: control group). After a month, the pre- and post-tests regarding text cohesion (e.g., reference item identification, function recognition of the reference test, reading comprehension test) were compared. The participants in the reference cohesion instruction condition had significantly greater scores than those in the control group. They concluded that implementing cohesion instruction and tasks was effective for improving the learners' knowledge of text cohesion.

Moreover, Safaie (2020) examined the effects of instruction on text connectives on 60 Iranian EFL learners. He explored the influences of implicit and explicit teaching of the linguistic devices that provide text cohesion in reading comprehension. The learners participated in either the explicit or implicit teaching groups: those in the former received an explanation for the connectives used in the texts, while those in the latter were only exposed to its examples. The participants' improvement in reading through the teaching

treatment was measured using a preliminary English test. The results of the pre- and post-English tests showed that both explicit and implicit teaching of cohesive connectives significantly affected the participants' reading performance. In addition, explicit teaching was found to be more effective than implicit instruction. Although this study was not conducted with the Japanese learners of English, it supported the importance of cohesion and the need for guidance for learners who are learning English as an EFL.

2.4 Summary of Previous Studies

Previous studies have revealed that revising text cohesion would be beneficial for those readers who have difficulty in connecting gaps within the text, and lack the ability to generate inferences, because text comprehension is gained through the readers' and texts' interactions (Gilabert et al., 2005; Linderholm et al., 2000; McNamara, 2001; McNamara et al., 1996; Ozuru et al., 2009; Vidal-Abarca et al., 2000). Text cohesion, which is defined as the extent of the text's ideas' explicitness, plays an important role in reading because it helps the readers to understand the text and construct a coherent mental representation. If the text cohesion is high, the readers can easily understand what the text is attempting to convey. This is because high-cohesion texts do not require the readers to speculate regarding the content that is not explicitly stated. However, if it is low, they must infer the necessary ideas to cover the gaps in the text (Hosoda, 2016; Kintsch, 1994; McNamara et al., 1996; Ozuru et al., 2010). Therefore, text cohesion is considered to be an essential factor affecting the readers' comprehension.

Previous research investigating the effect of revising text cohesion revealed that high-cohesion texts are helpful for the readers (e.g., Beck et al., 1991; Best et al., 2006; Gilabert et al., 2005; Hall et al., 2014; Linderholm et al., 2000; Loxterman et al., 1994; Oh, 2001; Ozuru et al., 2009, 2010; Vidal-Abarca et al., 2000; Yano et al., 1994) because

it reduces their inferences during reading. However, some studies have reported a reverse cohesion effect: Low-cohesion would be profitable for skilled readers (e.g., Hosoda, 2016; McNamara, 2001; McNamara et al., 1996; O'Reilly & McNamara, 2007). Nonetheless, other studies have debated that advanced readers take advantage of the high-cohesion text's information as compared to the low-cohesion one (Voss & Silfies, 1996). Hence, although previous research tested the influences of revising text cohesion, they did not reach an agreement about the effectiveness of text cohesion.

In addition, prior studies have tended to investigate the effects of manipulating text cohesion on the readers' comprehension, as measured by the post-reading or offline tasks (e.g., recall, open-ended questions, multiple-choice problems). However, the readers' online processes during reading have not been observed appropriately in past research (Hall et al., 2015). If text cohesion affects the readers' construction of mental representations, how and why it facilitates or impedes these representations needs to be investigated. To explore the impacts of text cohesion, describing the readers' processes in greater detail, using a think-aloud task as an online method, is needed (e.g., Horiba, 2013; Linderholm & van den Broek, 2002). Unlike offline tasks, it can help researchers to observe the readers' thoughts directly by making them verbalize their views.

Furthermore, previous L1 research has indicated that increasing text cohesion would be helpful for poor readers (e.g., McNamara et al., 1996; Ozuru et al., 2009); however, limited studies have investigated the role of revising it for Japanese L2 or EFL learners (Hosoda, 2016). As compared to the L1 readers, it has been reported that the EFL learners are poor at inferencing and have more problems in reading (Horiba, 1996). Thus, it is worth revealing the effect of revising text cohesion on the Japanese EFL learners by expanding the limitations of prior research (e.g., online methods). Moreover, applying this study's findings to support the L2 readers in terms of appropriate-cohesive text

assignments is needed. This is a feasible solution because the readers' abilities are not easily transformable.

2.5 Limitations of Previous Studies and the Purpose of the Current Research

As reviewed thus far, previous studies have revealed that text factors (e.g., text revision, text cohesion) may affect the reading and comprehension of the L1 and L2 learners (e.g., Hall et al., 2014; Hosoda, 2016; Vidal-Abarca et al., 2000; Yano et al., 1994). However, they have certain limitations and unexplored issues.

First, text cohesiveness in the English language materials that the Japanese learners of English are exposed to has not been comprehensively clarified. Although prior studies have shown that highly specialized and difficult texts are less cohesive, only a few have analyzed text cohesion in the materials for the Japanese EFL learners. Therefore, before examining the effects of cohesion revision on reading, it is primarily necessary to conduct research to elucidate the cohesiveness of the current materials used in Japan. Although the materials for young Japanese learners of English have been assessed (e.g., Nahatame, 2021), a large-scale text analysis has not been conducted thus far. To investigate the influences of text cohesion on reading comprehension, it is necessary to identify the current state of the materials and interpret the results accordingly.

Second, thus far, prior research has not reached a consensus on whether modifications that increase text cohesion work more effectively for reader comprehension, or conversely, whether low cohesion promotes comprehension. Some studies have revealed the beneficial effects of revising texts in terms of text cohesion on reading or comprehension, while others have demonstrated that it is not necessarily good for readers, especially those who inference appropriately during reading (McNamara, 2001; O'Reilly & McNamara, 2007). Thus, it is unclear whether text revision is effective for improving

L2 learners' reading comprehension. Since Japanese learners may not make good use of linguistic cues in texts (Ozono & Ito, 2003), it is important to clarify the impact of text cohesion on comprehension and how to apply it to instruction in the EFL context. However, few studies have investigated the effect of text revision on text cohesion in Japanese EFL learners' reading processes and comprehension.

Third, research examining text cohesion has focused on its impacts on comprehension in the L1 and L2 studies, as assessed using offline measures; additionally, they did not examine the processing of reading or whether readers evaluate the text cohesion. Therefore, whether they judge the changes in text cohesion and are affected by cohesion during reading, thereby changing their comprehension, remains unexplored.

Finally, text genre, a possible factor that affects the influences of text cohesion on reading, has not been investigated appropriately in previous studies. Furthermore, no prior research has examined whether it affects the effect of text cohesion on reading comprehension (to the knowledge of the authors, only the genre expectation studies have been conducted.) Previous research on historical manuscripts, scientific texts, and more have been limited to specific text genres and topics; they have been unable to provide a comprehensive and clear picture of whether the outcomes of text revision vary by genre. Considering that text cohesion affects the readers' comprehension, it is important to clarify the type of cohesion revision that is effective for a particular genre. This is an important topic for discussion.

Therefore, to address these issues, the current study aimed to elucidate the current state of text cohesion in the English language materials for the Japanese EFL learners (Study 1). Furthermore, it investigated the effects of manipulating text cohesion (high and low) on their judgments of text cohesion, comprehension, and processing, based on several empirical studies (Study 2). Study 1 comprised two surveys of text analysis, while

Study 2 consisted of six empirical studies with online (think-aloud task) and offline measures (e.g., written recall task). In addition, empirical research considered the influences of text genre and the readers' proficiency. The 13 research questions (RQs) of this study are summarized in Table 2.2.

Table 2.2

Research Questions and Focus of This Dissertation

S	TA / E	RQs	Object	Texts	Measure-ments
1	TA 1	**RQ1-1**: In English language textbooks for Japanese EFL learners, how do the texts for different grade levels differ in terms of cohesion?	text features	text-books	Coh-Metrix
	TA 2	**RQ1-2**: In the EIKEN tests for Japanese EFL learners, how do the texts for different grade levels differ in terms of cohesion?		EIKEN	
2	E1	**RQ2-1**: Are EFL readers aware of the degree of cohesiveness of a text? **RQ2-2**: Do EFL readers judge high- and low-cohesion texts differently in terms of their proficiency level?	awareness to cohesion	science topics	cohesion-judgment
	E2	**RQ3-1**: Does high or low cohesion affect Japanese EFL readers' surface-level comprehension of social expository texts? **RQ3-2**: Does high or low cohesion affect Japanese EFL readers' deep-level comprehension of social expository texts? **RQ3-3**: Does the cohesion effect on the comprehension of social expository texts vary as per the readers' proficiency level?	compre-hension	social topics	written recall, why-question
	E3	**RQ4-1**: Does high or low cohesion affect Japanese EFL readers' surface-level comprehension of scientific expository texts? **RQ4-2**: Does high or low cohesion affect Japanese EFL readers' deep-level comprehension of scientific expository texts? **RQ4-3**: Does the cohesion effect on the comprehension of scientific expository texts vary as per the readers' proficiency level?		science topics	
	E4	**RQ5-1**: Does cohesion affect inference generation among Japanese EFL readers while reading scientific expository texts? **RQ5-2**: Does the cohesion effect on inference generation from scientific expository texts vary as per the readers' proficiency? **RQ5-3**: Does the cohesion effect on inference generation from scientific expository texts vary depending on the readers' processing tendency during reading?	processes	science topics	think-aloud

Note. S = study; TA = text analysis; E = experiment; RQs = research questions; EFL = English as a Foreign Language, EIKEN = Jitsuyo Eigo Gino Kentei (Test in Practical English Proficiency).

Chapter 3

Study 1 Text Analysis on EFL Reading Materials

3.1 Text Analysis 1: Text Cohesion in Foreign Language Textbook for Japanese EFL Learners

3.1.1 Purpose and Research Questions

The purpose of Study 1 was to investigate the textual features of Japanese EFL reading materials to examine cohesiveness in the current English language materials before investigating the effects of text cohesiveness manipulation in the next study (Study 2).

Specifically, Study 1 consisted of two sub-investigations of text analyses (Text Analyses 1 and 2). Text Analysis 1 evaluated English textbooks for elementary, middle, and high school students; Text Analysis 2 examined texts in the reading section of the Jitsuyo Eigo Gino Kentei (Test in Practical English Proficiency), otherwise known as the EIKEN. Textbooks were chosen for evaluation because most young Japanese EFL readers learn English through compulsory education, and the EIKEN was chosen because it is a standardized test sponsored by the Ministry of Education, Culture, Sports, Science, and Technology in Japan. It is also one of the most common tests (i.e., over 3,500,000 learners took the EIKEN test in the 2020 fiscal year [FY]) used to measure learners' English skills in the Japanese EFL context.

Text Analyses 1 and 2 are based on studies sponsored by the Japan Textbook Research Center's "Grants for Graduate Student Monographs on Textbooks FY2018" and "EIKEN Grants" of EIKEN Foundation of Japan. The grant research was published in 2018 (Ogiso, 2018) and 2020 (Ogiso, 2020) and copyrighted by the Japan Textbook Research Center and the EIKEN Foundation of Japan, respectively.

The Coh-Metrix automatic text analysis tool (e.g., Graesser et al., 2014; McNamara et al., 2014) was used to clarify textual characteristics in the present study. This tool provided various perspectives for evaluating the lexical properties, syntactic features, and sentence cohesion of passages written in English. The tool's cohesion indexes include lexical redundancy between two sentences, statistically determined semantic relatedness of sentences, the usage frequency of conjunctions, causative verbs, verb tenses, phrase redundancy, and so on. The Coh-Metrix has been widely used in analyses of texts written for native English speakers (e.g., McNamara et al., 2014); more recently, it has also been applied to analyzing teaching materials for English learners (e.g., Nahatame, 2021; Nahatame & Kimura, 2019).In this study, the Coh-Metrix text analysis tool was used to analyze the primary text cohesion indicators—*reference cohesion* and *deep cohesion*—in educational materials commonly used for Japanese EFL learners. Using the above indexes, Text Analysis 1 was conducted to answer the following research question (RQ):

RQ1-1: In English language textbooks for Japanese EFL learners, how do the texts for different grade levels differ in terms of cohesion?

Previous studies have addressed text cohesion as one aspect of readability and reported that low-cohesion texts can make reading difficult even for primary readers in L1 (e.g., McNamara et al., 1996), suggesting that, in English, the higher the text level, the lower the text cohesion. Hence, this may also be true in EFL textbooks: the higher the text level of EFL textbooks and the intended students' levels, the lower the text cohesion.

3.1.2 Method

3.1.2.1 Materials

Analyzed texts

Text Analysis 1 focused on textbooks for junior high and high schools published by the three major textbook publishers in Japan: (i) Tokyo Shoseki, (ii) Kairyudo, and (iii) Sanseido. As for elementary schools, since foreign language (English) had not been implemented as a subject when this study was conducted, this study targeted only the teaching material series entitled "We can!" from which 20 textbooks were selected for analysis. Table 3.1 lists the textbooks and teaching materials used in Text Analysis 1.

Table 3.1

List of English Textbooks and Teaching Materials Analyzed in Text Analysis 1

School	Title (Publisher)
Elementary	"We can!" 1–2 (MEXT, 2018a, 2018b)
Middle	"New Crown" 1–3 (Negishi et al., 2016a, 2016b, 2016c), "Sunshine" 1–3 (Niisato et al., 2016a, 2016b, 2016c), "NEW HORIZON" 1–3 (Kasashima et al., 2019a, 2019b, 2019c)
High	"All Aboard!" (Kiyota et al., 2019a, 2019b, 2019c), "PROMINENCE English Communication I–III" (Tanabe et al., 2019a, 2019b, 2019c), "New Discovery English Communication I–III" (Namai et al., 2019a, 2019b, 2019c)

Note. High school students' textbooks were those of Communicative English I–III.

Among the passages in the target textbooks or teaching materials, the English passages appearing in the Lesson/Unit/Program (hereinafter called *unit*) were used for analysis;

however, all alphabet-only lessons and texts designed for speaking or writing activities were excluded from the analysis of text cohesion. Furthermore, the need to include dialogs was debatable; however, since the regular lessons in many textbooks included dialogs (designed for not only speaking but also reading or grammar teaching), they were included in our analysis. Supplementary readings and texts for extensive reading were included as well. Finally, two elementary school, 107 middle school, and 113 high school units in the textbooks were used for analysis.

Analysis tool

An automated text-evaluation tool, Coh-Metrix, was employed to discern the text features of the various test materials. It analyzes texts at multiple levels of language and discourse, ranging from words to discourse genres (e.g., Graesser et al., 2011; Graesser et al., 2004). More details are available in McNamara et al. (2014), and the wide range of language and discourse measures provided by this computational tool can be found on the Coh-Metrix website (http://tea.cohmetrix.com/).

Five indexes in the Coh-Metrix Text Easability Assessor considered sufficient to grasp the general text features of each reading text were adopted for this study: *narrativity*, *syntactic simplicity*, *word concreteness*, *reference cohesion*, and *deep cohesion*. These indexes were chosen because they are simple yet good measures for identifying the general features of the text, and the definitions of each are presented in Table 3.2. The text was analyzed, and each index feature was rated on a scale from 0, *very difficult* to 100, *very easy*. In addition to these indexes, since the analysis was based on English texts from different school types, a difficulty index, Flesch-Kincaid Grade Level, which evaluates the difficulty of English texts based on the length of words in the text, was also calculated.

Table 3.2

Indexes and Definitions of Text Analysis Based on McNamara et al. (2014)

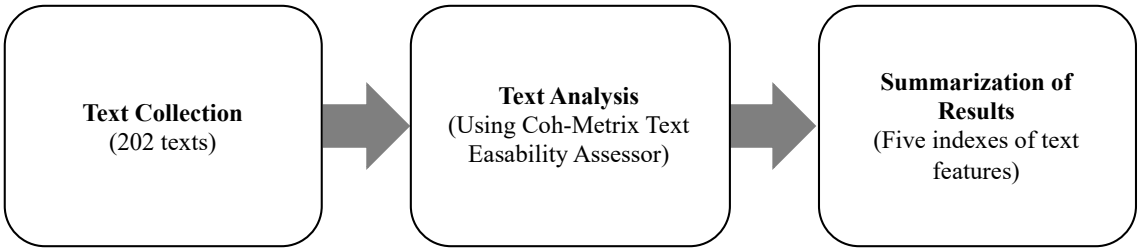
Index	Definition
Narrativity	The degree to which a text contains narrative elements and is familiar
Syntactic Simplicity	The degree to which a text contains fewer and simpler words
Word Concreteness	The degree to which a text contains imaginary and concrete words
Reference Cohesion	The degree to which a text is explicit with overlapping elements
Deep Cohesion	The degree to which a text contains logical and causal relations

3.1.2.2 Procedure

This text analysis was conducted by the researcher alone. Figure 3.1 shows the procedure for Text Analysis 1.

Figure 3.1

Procedure of Text Analysis 1



First, English passages (i.e., narrative, explanatory, and essay texts) included in the textbooks or teaching materials were converted into digital text data using a scanner. In

instances of misread paragraphs and misrecognized characters, the researcher manually fixed the digital data. Second, 222 units of English text were analyzed using the analysis tool. If there were several small parts in a lesson (unit, program), they were summed and analyzed as a whole text. Finally, five textual characteristics and difficulty indexes were summed.

3.1.2.3 Analysis

First, the textual characteristics of the English content in the textbooks and teaching materials were calculated using the Coh-Metrix. Next, for each calculated value, the averages by school type (i.e., elementary, middle, and high school) were summed; then, the calculated mean values of the characteristics and their changes by grades or difficulty of the textbooks were confirmed. Since previous studies have also interpreted trends using mean values rather than statistical analysis (Nahatame, 2017; Mori, 2018), this study adopts the same procedure. In addition, due to the extreme differences in the number of texts per school level, statistical comparisons were not made.

3.1.3 Results

Table 3.3 and Figure 3.2 show the mean values and standard deviations of English textual characteristics of English textbooks or teaching materials used in elementary to high school.

Table 3.3

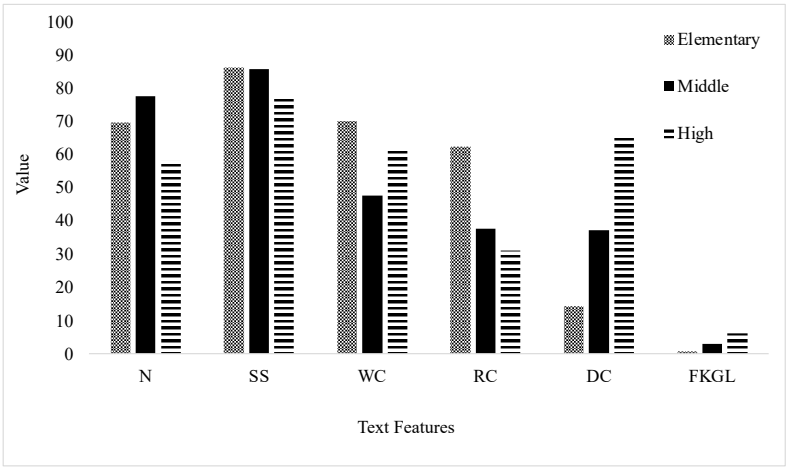
Descriptive Statistics of English Text Characteristics and Readability Indexes in Major English Teaching Materials and Textbooks by School Type in Text Analysis 1

School	Elementary (n = 2)		Middle (n = 107)		High (n = 113)	
Textual Characteristics	*M*	*SD*	*M*	*SD*	*M*	*SD*
Narrativity	69.72	10.24	77.61	15.12	58.05	21.29
Syntactic Simplicity	86.12	11.26	85.84	10.43	76.92	11.59
Word Concreteness	70.09	6.03	47.71	28.84	61.89	23.52
Reference Cohesion	62.31	32.85	37.60	19.74	31.02	18.36
Deep Cohesion	14.42	10.41	37.27	23.59	65.65	22.35
FKGL	0.80	1.13	3.10	1.06	6.23	1.58

Note. n = targeted unit of textbooks; FKGL = Flesch-Kincaid Grade Level. Indexes except Flesch-Kincaid Grade Level have a value between 0 and 100.

Figure 3.2

Transitions of English Text Characteristics and Readability Indexes in Major English Teaching Materials and Textbooks by School Type in Text Analysis 1



Note. N = narrativity; S = syntactic simplicity; WC = word concreteness; RC = reference cohesion; DC = deep cohesion; FKGL = Flesch-Kincaid Grade Level. Indexes except Flesch-Kincaid Grade Level have a value between 0 and 100.

First, let us look at the linguistic features calculated using the Coh-Metrix. The results of the Flesch-Kincaid Grade Level, a traditional difficulty index, confirmed that English sentences became more difficult as the school level increased. The value of syntactic simplicity, which indicates how simple the syntactic structure of English sentences is, decreased as the school type increased. This indicates that English sentences in textbooks are gradually becoming more complex. As for Narrativity and Word Concreteness, the values were not necessarily linear to the level of the school.

Regarding text cohesion indexes, the reference cohesion value, which may interfere with text comprehension when it is low, became lower as the school level increased: M = 62.31, SD = 32.85 for teaching materials for elementary school, M = 37.60, SD = 19.74 for textbooks for middle school, and M = 31.02, SD = 18.36, for high school textbooks.

Meanwhile, for deep cohesion, which reflects causal and intentional connectives in texts, the value increased as the school (grades) level increased: M = 14.42, SD = 10.41 for teaching materials for elementary school, M = 37.27, SD = 23.59 for middle school textbooks, and M = 65.65, SD = 22.35 for textbooks for high school textbooks.

3.1.4 Discussion

Response to RQ1-1: ***In English language textbooks for Japanese EFL learners, how do the texts for different grade levels differ in terms of cohesion?***

The text analysis results showed that the texts in English language textbooks for Japanese EFL learners differed in both reference and deep cohesion depending on the grade levels. The more difficult the English texts, the lower the reference cohesion. As for deep cohesion, which is reflected by the number of causal and intentional conjunctions, it was found to be most highly cohesive in the textbooks for higher grade levels.

This result of reference cohesion supports the results also claimed in reading

comprehension studies. The more difficult the text is, the less explicit are the relationships among the elements in the text, and the more the reader has to guess the relationships (e.g., McNamara & Kintsch, 1996). Considering that reference cohesion is calculated by the degree to which a text is explicit and uses overlapping elements, this result suggests that English textbooks for Japanese EFL readers gradually become less cohesive and contain fewer overlapping elements. Therefore, if materials are not suitable for the reader's level, there will be a mismatch between their reading comprehension skill and the explicit nature of the text. This mismatch can lead to reading difficulties.

However, in contrast to the reference cohesion, deep cohesion increased as the text's difficulty increased. This finding may seem surprising, but it may be explained by the fact that low cohesion is not necessarily linked to difficulty. Nahatame (2021) analyzed EFL materials for young children and found that the cohesion factor did not completely explain the difficulty of English texts. Moreover, deep cohesion is determined by the degree to which the text contains causal and intentional conjunctive expressions; thus, it may be reasonable to find that the difficulty index and the number of conjunctions do not correlate well in texts that are designed to be difficult due to other lexical and grammatical factors.

In addition, this text analysis included various kinds of texts, from very short texts used at the elementary school level to dialog texts in which multiple characters were talking. In the case of other text types, deep cohesion may decrease as the level of text difficulty increases, while, in the case of academic texts, deep cohesion may be higher even if the English text is more advanced because of the nature (and necessity) of explaining causality in academic texts (e.g., thesis, journals). Lightman et al. (2007) revealed that texts in different genres (science, history, and narrative texts) had different degrees of cohesion. They suggested that the cohesion level of the science texts is higher

than that of the historical and narrative texts. Thus, this study's results may be due to the various kinds of texts included in the analysis.

Although determining the relationship between the level of difficulty and deep cohesion may be problematic based only on the materials analyzed in this study, the results certainly suggest that the degree of text cohesion differs depending on the grade/difficulty level in regard to textbooks in terms of both reference and deep cohesion.

However, this text analysis was conducted before 2020, and the analyzed textbooks did not follow the latest government course guidelines, which included English as a compulsory subject at the elementary level. Thus, new elementary school English textbooks as a compulsory subject were not included in this analysis. Future studies should analyze the newest textbooks for Japanese EFL learners to gain insights about the latest versions of textbooks from elementary to high school level.

3.1.5 Conclusion of Text Analysis 1

Text Analysis 1 aimed to investigate the textual features of Japanese EFL reading materials (English teaching materials and textbooks used in Japan) using a Coh-Metrix text analyzer. In Text Analysis 1, reference cohesion and deep cohesion, which are indicators of the explicitness and connectedness of English sentences (i.e., text cohesion), were calculated for English textbooks from elementary to high school, and the changes in the values related to the difficulty level were observed.

The results showed that reference cohesion decreased as the difficulty level increased, suggesting that it may be hard for language learners to understand the reference cohesion—that is, the correspondence between pronouns—if the material is too difficult as compared to their reading level. This finding is especially notable for designers of educational materials. In contrast, for deep cohesion, the higher the grade of the material,

the higher the cohesive value. Although this finding is interesting, it may not be consistent with other types of textbooks in terms of the number of words and vocabulary used, given that school textbooks are designed with the developmental stage of the reader in mind.

Conducting such an analysis on a large sample of textbooks would be valuable in EFL reading research, as it was difficult to generalize the results of this study because only a few materials were included. For example, Text Analysis 1 was limited to the analysis of 20 textbooks (222 passages). Therefore, to gain a better understanding of more general trends, a comprehensive survey of certified textbooks is necessary. It is also recommended that the characteristics and difficulty level of each textbook in a larger study be clarified in detail. The results of such a study, if disclosed to the creators of teaching materials and teachers, may contribute to advances in teaching and the use of materials with more levels of difficulty.

Therefore, the aim of Text Analysis 2 was to examine the reference and deep cohesion and other text characteristic indexes, with respect to the difficulty level in different grades, in texts from large-scale English language tests.

3.2 Text Analysis 2: Text Cohesion in Foreign Language Test Passages for Japanese EFL Learners

3.2.1 Purpose and Research Questions

Text Analysis 2 was conducted with the same motive as Text Analysis 1: To explore the textual features of Japanese EFL reading materials by analyzing texts used in the large-scale standardized English test (EIKEN) for Japanese EFL learners. Throughout this chapter, Study 1 attempts to explore whether materials for Japanese EFL readers vary in their text cohesion indexes before conducting the experimental study, which investigates the effects of text cohesion manipulation on Japanese EFL readers. Text Analysis 1 was conducted to observe changes in cohesion in different-level English texts and obtain evidence to support the hypothesis that EFL readers of different proficiency levels must be able to comprehend English texts with various degrees of cohesion.

To expand the findings of Text Analysis 1, in Text Analysis 2, the EIKEN texts were analyzed. There were three main reasons to use the EIKEN texts. First, the EIKEN is a standardized test sponsored by the Ministry of Education, Culture, Sports, Science, and Technology and is a common test for measuring learners' English skills in the Japanese EFL context. Second, compared to the textbooks analyzed in Text Analysis 1, the EIKEN texts contain more expository texts, which have been used for cohesion manipulation in previous L1 studies. Third, the EIKEN test has seven grade levels (i.e., Grades 5, 4, 3, Pre-2, 2, Pre-1, and 1) and is based on the knowledge gained from creating and administering tests over the past 50 years. Thus, Text Analysis 2 was designed to explore expository texts used in the past EIKEN tests, and the following RQ is presented.

RQ1-2: In the EIKEN tests for Japanese EFL learners, how do the texts for different grade levels differ in terms of cohesion?

To answer the RQ in Text Analysis 2, this study applied the Coh-Metrix (consistent with Text Analysis 1). It seemed likely that similar to Text Analysis 1, there would be less reference cohesion, which reflects the degree of explicitness and overlap in the text as the difficulty of the text increases. Moreover, the results for deep cohesion were not expected to be the same as those for Text Analysis 1 since they depend on whether the text type includes many conjunctions, and it is difficult to predict deep cohesion for low-level texts because such easy texts are inherently limited in terms of their content and length (the number of words). However, for some grade levels of reading (e.g., above 2nd grade), deep cohesion may increase as the text difficulty increases.

3.2.2 Method

3.2.2.1 Materials

Analyzed texts

The EIKEN Test in Practical English Proficiency (EIKEN) (constructed by the EIKEN Foundation, Japan) was adopted to analyze reading materials of different difficulty levels. As mentioned in Section 3.2.1, the EIKEN test was used because a) it is a standardized test widely used in Japan, b) it has many expository texts in the reading section, and c) it has seven grade levels (i.e., Grades 5, 4, 3, Pre-2, 2, Pre-1, and 1). See Table 3.4 for reference to EIKEN grades and their relation to CEFR English level.

Table 3.4

EIKEN Grades and the Equivalent Ability Description

EIKEN Grade	Example of Recognition/Use	CEFR Level
1	International admissions to graduate and undergraduate programs	C1
Pre-1		B2
2	MEXT benchmarks for high school graduates	B1
Pre-2		A2
3	MEXT benchmarks for junior high school graduates	A1

Note. CEFR = Common European Framework of Reference for Languages; MEXT = The Japanese Ministry of Education, Culture, Sports, Science, and Technology. This table was adopted and created based on the information on the EIKEN official website (https://www.eiken.or.jp/eiken/en/grades/).

This study analyzed 501 texts retrieved from the reading section of previous EIKEN tests. Specifically, data from 1998 to 2016 for five levels, from Grades 3 to 1, were adopted. Most of the 501 texts were expository texts or essays with varied themes including science, music, and philosophy. Note that texts of e-mails, notice boards, and texts with blanks were excluded from the analysis. The average number of words and readability of the texts analyzed in the pilot study are shown in Table 3.5.

Table 3.5

Words and Readability of Texts Analyzed in Text Analysis 2

	1 (n = 160)		Pre-1 (n = 132)		2 (n = 97)		Pre-2 (n = 66)		3 (n = 46)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Words	547.19	147.93	414.52	73.31	354.73	19.96	291.65	18.97	253.13	12.15
FKGL	12.95	1.90	12.11	1.37	9.33	1.16	7.89	0.99	5.88	1.01

Note. n = number of texts analyzed; *SD* = standard deviation; FKGL = Flesch-Kincaid grade level.

Analysis tool

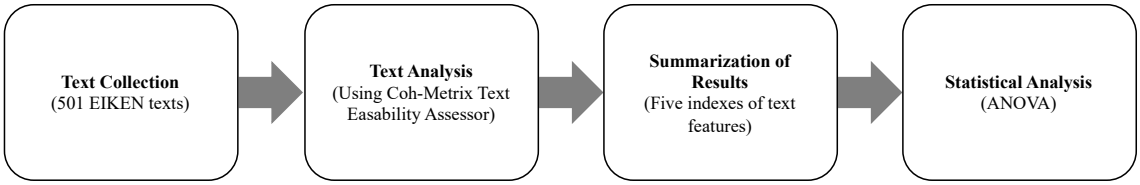
To obtain the text features of different-level test materials, the Coh-Metrix was employed. To check the general text features of each reading text, five indexes in the Coh-Metrix Text Easability Assessor were used: *narrativity*, *syntactic simplicity*, *word concreteness*, *reference cohesion*, and *deep cohesion*. These indexes were chosen because they are simple but good measures for assessing the general features of the text and they include measurement of both kinds of text cohesion (reference and deep). The definitions of each index are presented in Table 3.2 in Section 3.1.2.1.

3.2.2.2 Procedure

This text analysis was conducted by the researcher alone. The procedure for Text Analysis 2 is shown in Figure 3.3.

Figure 3.3

Procedure of Text Analysis 2



First, 501 reading texts were collected from the reading section of past EIKEN tests. The EIKEN Foundation was notified of this collection before the study. After receiving approval from the foundation to use their data, the EIKEN texts were converted into digital text data using a scanner. Any misread paragraphs and misrecognized characters were manually corrected by the researcher.

Then, each text was analyzed using the Coh-Metrix Text Easability Assessor (i.e., online text analyzer), and the basic data of each passage (e.g., words and readability) were checked. Next, the researcher summarized the results of the text analyses in terms of the five text features: narrativity, syntactic simplicity, word concreteness, reference cohesion, and deep cohesion. Finally, the average scores of the five indexes were tested using statistical analysis to check the differences between text features of various difficulty levels.

3.2.2.3 Analysis

The textual characteristics of the English texts were calculated using the Coh-Metrix, and their calculated values and descriptive statistics were summed using EIKEN grade levels. Next, to confirm any differences between reading texts of different difficulty levels, the average scores were calculated for each of the five text feature indexes, and a

repeated two-way analysis of variance (ANOVA) was conducted to examine the relationships between the difficulty levels and text features. The significance level of the analysis was set to $\alpha < .05$.

3.2.3 Results

Table 3.6 summarizes the mean scores and standard deviations for the texts for each grade in terms of the five analysis indexes: narrativity, syntactic simplicity, word concreteness, reference cohesion, and deep cohesion. In general, their scores were lower at the more difficult levels (i.e., Grade 1) than at the easier levels (i.e., Grade 3). That is, the easier the text level, the higher the indexes for easiness. Considering that each index reflects the familiarity, simplicity, and concreteness of words, the results suggest that the lower-level texts contain words that are easily readable.

With regard to text cohesion, which is the focus of this study, the related indexes (i.e., reference cohesion and deep cohesion) were lower for higher-level texts than for lower-level texts. This result suggests that higher-level texts are less cohesive as compared to the lower-level texts (reference cohesion: $M = 16.74$, $SD = 14.26$ for Grade 1, $M = 19.73$, $SD = 16.32$ for Grade Pre-1, $M = 44.95$, $SD = 20.76$ for Grade 2, $M = 56.74$, $SD = 19.98$ for Grade Pre-2, $M = 68.22$, $SD = 18.50$ for Grade 3; deep cohesion: $M = 63.75$, $SD = 18.94$ for Grade 1, $M = 64.16$, $SD = 21.60$ for Grade Pre-1, $M = 74.64$, $SD = 19.53$ for Grade 2, $M = 77.97$, $SD = 20.75$ for Grade Pre-2, $M = 66.72$, $SD = 23.58$ for Grade 3).

Table 3.6

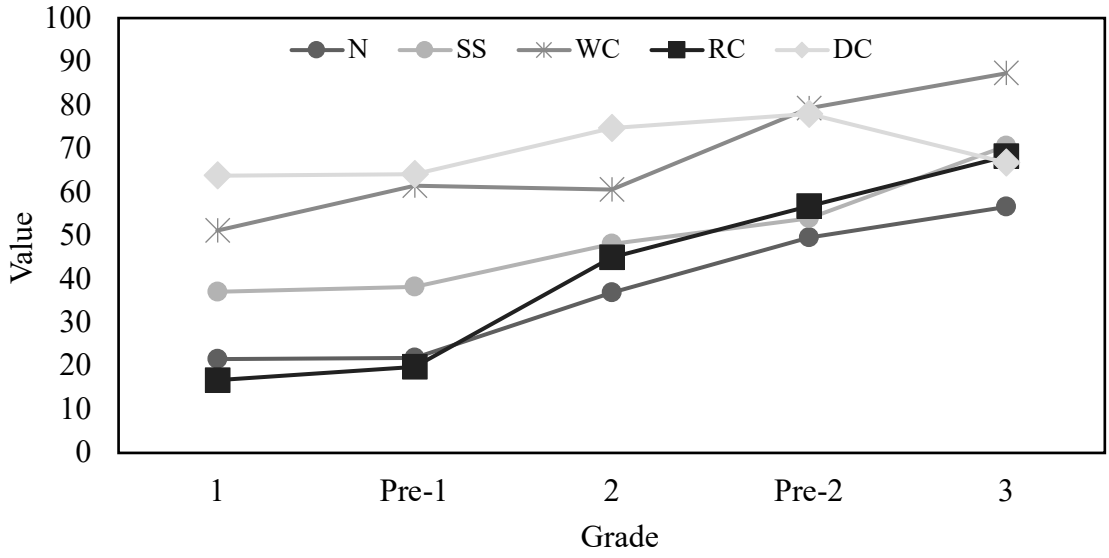
Means and SDs of Five Text Features Calculated by Coh-Metrix in Text Analysis 2

	1		Pre-1		2		Pre-2		3	
	(*n* = 160)		(*n* = 132)		(*n* = 97)		(*n* = 66)		(*n* = 46)	
TF	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
N	21.58	12.13	21.83	9.60	36.86	13.46	49.53	15.55	56.57	18.24
SS	37.06	14.53	38.21	14.16	48.10	14.43	53.89	15.48	70.59	13.12
WC	51.13	20.49	61.45	22.22	60.63	23.78	79.23	16.51	87.30	11.76
RC	16.74	14.26	19.73	16.32	44.95	20.76	56.74	19.98	68.22	18.50
DC	63.75	18.94	64.16	21.60	74.64	19.53	77.97	20.75	66.72	23.58

Note. n = number of texts analyzed; TF = text feature; N = narrativity; SS = syntactic simplicity; WC = word concreteness; RC = reference cohesion; DC = deep cohesion.

Figure 3.4

Text Features by Each Grade Calculated by Coh-Metrix in Text Analysis 2



Note. N = narrativity; SS = syntactic simplicity; WC = word concreteness; RC = reference cohesion; DC = deep cohesion.

To examine the relationship between difficulty levels and text features, a 5 (text feature: narrativity, syntactic simplicity, word concreteness, reference cohesion, and deep cohesion) × 5 (grade: 1, Pre-1, 2, Pre-2, 3) two-way ANOVA was conducted on the mean scores of Coh-Metrix Text Easabilities for the past EIKEN texts, with text features as a within-participants variable and grade as a between-participants variable. Note that the presumption of statistics, such as normality, was confirmed before the analysis. The results are summarized in Table 3.7.

Table 3.7

ANOVA Results for Main Effects and Interaction Effects of Text Feature and Grade Level on Five Indexes of Text Analysis 2

Source	*SS*	*df*	*MS*	*F*	*p*	η^2
			Between Subjects			
Grade (G)	312221.65	4.00	78055.41	228.861	<.001***	.21
Error	169165.91	496.00	341.06			
			Within Subjects			
Text Feature (TF)	364800.76	3.42	106814.42	305.47	<.001***	.24
TF × G	78762.53	13.66	5765.46	16.49	<.001***	.05
Error (TF)	592334.06	1693.98	349.67			
Total	1517284.92	2211.05				

Note. ANOVA = analysis of variance. ****p* <.001.

The results showed that there were significant main effects for text features with small effect size, $F(3.42, 1) = 305.47$, $p < .001$, $\eta^2 = .24$, grade with small effect size, $F(4, 496) = 228.86$, $p < .001$, $\eta^2 = .21$, and an interaction effect for Text Feature × Grade with medium effect size, $F(13.66, 4) = 16.49$, $p < .001$, $\eta^2 = .05$.

Because this study focused on exploring the differences among text features in different-level texts, the following results were focused on the simple main effects of text features. The results of multiple comparisons with Bonferroni correction showed that the simple main effects of all text features were statistically significant: $F(4, 496) = 119.12$, $p < .001$ for narrativity, $F(4, 496) = 62.40$, $p < .001$ for syntactic simplicity, $F(4, 496) = 39.82$, $p < .001$ for word concreteness, $F(4, 496) = 142.38$, $p < .001$ for reference cohesion; and $F(4, 496) = 9.30$, $p < .001$ for deep cohesion.

The following multiple comparisons with Bonferroni correction tested the difference between 10 combinations of two out of five grades for each text feature (i.e., narrativity, syntactic simplicity, word concreteness, reference cohesion, and deep cohesion). First, for narrativity, there were statistically significant differences between all eight combinations, except two parts of Grade 1–Pre-1 ($p = 1.00$) and Pre-2–3 ($p = .052$) in narrativity (ps $< .001$).

As for syntactic simplicity, significant differences between all eight combinations of text features were confirmed (ps $< .001$), except for two parts of Grade 1 vs. Pre-1 ($p = 1.00$) and 2 vs. Pre-2 ($p = .160$). With regard to word concreteness, all seven combinations were significant (ps $< .001$), except for three pairs of Grade 1 vs. 2 ($p = .004$), Pre-1 vs. 2 ($p = 1.00$), and Pre-2 vs. 3 ($p = .422$).

Regarding reference cohesion, which was one of the text cohesion indexes and a particular target in the present study, all combinations, except one pair of Grade 1 vs. Pre-1 ($p = 1.00$), were significantly different (eight pairs, ps $< .001$; one pair of Grade Pre-2

vs. 3, p = .007). Lastly, in the case of deep cohesion, five combinations of grades [Grade 1 vs. 2, 1 vs. Pre-2, Pre-1 vs. 2, & Pre-1 vs. Pre-2 (ps <.001), Pre-2 vs. 3 (p = .046)] were significantly different; five pairs [Grade 1 vs. Pre-1 (p = 1.00), 1 vs. 3 (p = 1.00), Pre-1 vs. 3 (p = 1.00), 2 vs. Pre-2 (p = 1.00), 2 vs. 3 (p = .320)] were not significantly different.

In summary, the results indicate that the difficulty indexes of the texts used in the past EIKEN tests display a variety in terms of words (e.g., simplicity) and cohesion (e.g., reference cohesion, deep cohesion). Furthermore, although not all pairs of different grades were significantly different, a certain tendency toward a difference between high-grade and low-grade texts was confirmed through analysis: more difficult texts involve unfamiliar topics (i.e., narrativity), more complicated syntactic organization (i.e., syntactic simplicity), harder-to-imagine words (i.e., word concreteness), and lower text cohesion (i.e., reference cohesion) and coherence (i.e., deep cohesion).

As for reference cohesion, nine of the ten pairs of different grades were significantly different. In other words, text cohesion was significantly different in different-level texts, but deep cohesion was not significantly different from text cohesion. As expected, there were significant differences in text coherence between different-level texts, but adjacent-level texts (e.g., Grade 1 vs. Pre-1) were not very different. The results, when taken together, indicate an overall tendency for the five text features calculated by the Coh-Metrix to differ significantly between texts for different grades. However, the extent of the differences is dependent on text features.

3.2.4 Discussion

Response to RQ1-2: ***In the EIKEN tests for Japanese EFL learners, how do the texts for different grade levels differ in terms of cohesion?***

The result of the ANOVA conducted on the 501 EIKEN text features showed that reference cohesion significantly differed between most pairs of different grades, and deep cohesion also significantly differed between some pairs of texts for different levels. As shown in Table 3.6 and Figure 3.4, the mean scores for reference cohesion gradually decrease as the grade level of the text increases. This result suggests that higher-level texts are less cohesive than lower-level texts.

This result was consistent with previous L1 studies (e.g., McNamara & Kintsch, 1996) that have asserted that less cohesive texts are difficult for readers and reference cohesion. Disconnected (or less cohesive) information, which does not have enough explanation, has been shown to impede readers' comprehension (O'Reilly & McNamara, 2007), and less-skilled readers are not as good at generating inferences to connect information or at guessing the implicit message of the text (Linderholm et al., 2000; McNamara & Kintsch, 1996). Hence, it is reasonable that the results confirm that more difficult texts are less cohesive and easier lower-level texts are more cohesive.

However, this result is inconsistent with that of McNamara et al.'s (2012) study. They analyzed English corpus and found that referential cohesion increased across the grade levels; furthermore, they considered this to be a counterintuitive result due to the short sentences in the lower-grade texts. Additionally, they suggested that educators should provide scaffolding for unskilled readers to address the gaps in the low-cohesion texts. Given that the current study was conducted using EFL materials in which the publisher edited and rephrased the text elements (vocabulary, grammar), the inconsistency between L1 and EFL text analysis results is understandable.

Moreover, this study revealed that the reading materials for EFL learners have similar characteristics to those for L1 readers. Although the types of materials analyzed in Text Analyses 1 and 2 were different, the reference cohesion moved counter-proportionally to the difficulty and grade level of the English text. Furthermore, this finding indicates that proficient readers, who are expected to understand higher-level texts, should be able to comprehend low-cohesion texts.

The result of the ANOVA also showed that deep cohesion was significantly different in half of the pairs consisting of texts from two different grades. As shown in Table 3.6 and Figure 3.4, the mean scores for deep cohesion gradually decrease as the grade of the text increase such that the medium grade texts (e.g., for Grade 2, Pre-2) were relatively highly cohesive as compared to those for higher and lower grades (e.g., Grades 1 and 3). This result suggests that the deep cohesion of texts was different in texts of different grade levels. However, the manner of change between grades was not linear, as with text cohesion.

This finding was interesting in that the manner of change was quite different from that of the reference cohesion. Whereas reference cohesion decreased when the text level increased, deep cohesion did not demonstrate such a tendency. Rather, deep cohesion was highest in the intermediate-level text in Text Analysis 2. This result is inconsistent with the findings of Text Analysis 1, and, given that this study represents a pioneering attempt to analyze and capture EFL text features, there are not enough criteria or studies available for comparison at this time. Nevertheless, it suggests that the transition of deep cohesion in English texts does not seem to be as simple as that of reference cohesion.

Furthermore, there is a possible interpretation presented in this paper. That is, the words or information contained in the text affect its deep cohesion. In the original EIKEN texts, lower-grade texts were limited in word length (i.e., the mean total of words in the

text was 253.13 in Grade 3) because longer texts are regarded as too difficult to be understood by the intended examinees in Grade 3. However, the intermediate-level texts had vast amounts of information (i.e., the mean total of words in the texts was 357.73 in Grade 2). Thus, a sufficient number of connections for causal coherence may exist in texts containing a certain number of words (e.g., McNamara et al., 2012).

On the contrary, at the most difficult level, excessive information was included in the texts (i.e., the mean total number of words in the text was 547.19 in Grade 1 of the EIKEN texts). Grade 1 texts had a substantial volume of information, which might have reduced the possibility of including connectives to make room for other important contents or difficult words.

However, the mean value for deep cohesion in the analysis did not change significantly as compared to the reference cohesion index (see Figure 3.4). Given that this analysis did not discern and include factors such as text topics, genres, and words, these results should be interpreted carefully.

According to Warren (2012), text coherence is defined as the way in which text elements, such as sentences or phrases, belong together in terms of their meaning. Because deep cohesion is a linguistic indicator of semantic and causal connections, and it may be considered a concept similar to text coherence. In previous studies, text cohesion and coherence have often been considered similar concepts. Whereas text cohesion contributes to textual coherence, it is not always necessary for text coherence. Thus, a highly cohesive text may not be sufficiently coherent (Traxler, 2011). In contrast, a low-cohesive text is not always low-coherent. Assuming that deep cohesion is different from coherence in terms of textual characteristics and considering such complicated relations between the amount, cohesion, and coherence of text, it is reasonable to expect the chevron pattern in Figure 3.4. A non-linear change in deep coherence displayed a non-

from lower- to higher-grade texts.

3.2.5 Conclusion of Text Analysis 2

Text Analysis 2 was conducted to reveal the characteristics of the texts aimed at Japanese EFL learners. Specifically, this study aimed to explore whether materials of different difficulty have varied text features from the standpoint of text cohesion.

A total of 501 texts used in past EIKEN tests were analyzed using the Coh-Metrix Text Easability Assessor. To observe text cohesion, reference cohesion was calculated and analyzed. Reference cohesion was defined as the degree to which a text is explicit with overlapping elements, and deep cohesion, defined as the degree to which a text contains logical and causal relations.

The results showed that reference cohesion was mostly significantly different between different-level texts. Reference cohesion was higher in lower-level (i.e., easy) texts; text cohesion was lower in higher-level (i.e., more difficult) texts. For deep cohesion, it was sometimes significantly different between different-level texts. However, text coherence changed in a chevron pattern: intermediate-level texts had the highest text coherence. Based on the results of this study, the hypothesis is supported—the readers with different levels of EFL proficiency must understand texts with different levels of cohesiveness.

3.3 Conclusion of Study 1

Study 1 was composed of two text analyses (Text Analysis 1 and Text Analysis 2) exploring the textual features (text cohesion) of Japanese EFL reading materials. This study was necessary to clarify cohesiveness in the current English language materials to conduct Study 2, an investigation of the effects on Japanese EFL learners when the

cohesiveness of the text is manipulated (Study 2).

Study 1 set two RQs: *In English language textbooks for Japanese EFL learners, how do the texts for different grade levels differ in terms of cohesion?* (RQ1-1) and *In the EIKEN tests for Japanese EFL learners, how do the texts for different grade levels differ in terms of cohesion?* (RQ1-2). To address each RQ, Text Analysis 1 was conducted to examine textbooks, and Text Analysis 2 was conducted to analyze large-scale English tests.

The results of the Text Analysis 1 revealed that reference cohesion in English textbooks becomes less cohesive as the school grade increases. However, the results for deep cohesion were not necessarily linear, and it was found that high school textbooks, in which the text length increased, had the deepest cohesion. The results of Text Analysis 2 with 501 texts from past EIKEN tests by the Coh-Metrix suggested that reference cohesion was used less as the text became more difficult. Moreover, for deep cohesion, cohesion tended to decrease as the text became more difficult (from Grade Pre-2 to Grade 1). However, it was confirmed that cohesion was low even in Grade 3, the easiest grade of the texts analyzed.

Reference cohesion was found to be inversely proportional to the difficulty of the English texts in both investigations while deep cohesion results were not consistent, but the transition was not necessarily the same as for reference cohesion. These results indicate that the relationship between text cohesion and difficulty can change depending on the type of text cohesion and material.

However, reference cohesion, which indicates the overlapping of content words and explicitness of information, was related to difficulty in both English textbooks and tests for EFL learners, indicating that revising texts in relation to reference cohesion could affect their difficulty and readers' comprehension. Moreover, the results suggest that

appropriate cohesion is necessary in English texts for readers to understand the information; that is, EFL materials for younger and less proficient readers should be highly cohesive, although this is not necessary for skilled readers. However, as indicated by McNamara et al. (2012), when the text cohesion is excessively low for unskilled readers, scaffolding should be provided to help them.

Although there were some biases in the materials and topics analyzed in Study 1, further studies are needed for comparison and to make a general conclusion. Based on the relationship between difficulty and text cohesion in Japanese English teaching materials, Study 2 involved several empirical studies investigating whether manipulating the text cohesion of English texts affects the awareness of cohesion, processing during reading, and text comprehension of Japanese readers who are learning English.

Chapter 4

Study 2 Effects of Text Cohesion on Japanese EFL Reading

4.1 Experiment 1: Effects of Text Cohesion on EFL Readers' Judgment

4.1.1 Purpose and Research Questions

The purpose of Study 2 was to examine the effects of manipulating text cohesion (high and low) on Japanese EFL learners' reading. As mentioned earlier, Study 2 consisted of four experimental studies, Experiments 1, 2, 3, and 4. In addition, Experiments 2 and 3 dealt with learners at different proficiency levels, including follow-up studies (Experiments 2-A, 2-B, 3-A, and 3-B). Refer to the subsequent chapters for details of Experiments 2, 3, and 4.

Experiment 1 was designed to investigate readers' monitoring or judgment of text characteristics (i.e., text cohesion) based on the results of the text analyses in Study 1. The results of Study 1 suggested that readers at different proficiency levels have to adapt to different texts that have varying characteristics (i.e., the degree of text cohesion). When reading higher-level texts with lower cohesion, readers need to infer the implicit ideas described in them. In contrast, reading lower-level texts with higher cohesion requires readers to understand the explicit relations between sentences. These textual features in subject matter with varying difficulty were confirmed in Text Analyses 1 and 2 analyzing English texts for Japanese EFL readers, and it suggested that EFL readers need to understand different degrees of text cohesion depending on their English proficiency level.

However, whether EFL readers are aware of text cohesion has not yet been researched. It might be useful to discern text cohesion for reading text strategically, and performing additional reading behavior (e.g., lookback, bridging information) to understand it coherently. Therefore, this experiment investigated whether Japanese EFL

readers, with different reading proficiency levels, could be aware of texts with varying degrees of cohesion. The two research questions in Experiment 1 were as follows:

RQ2-1: Are EFL readers aware of the degree of cohesiveness of a text?

RQ2-2: Do EFL readers judge high- and low-cohesion texts differently in terms of their proficiency level?

To measure readers' evaluation of text cohesion, this study conducted a text cohesion judgment task (e.g., Helder et al., 2016), assessing their sensitivity toward the same. In the task, readers were asked to evaluate whether the text was explicit and clearly described or not. In this study, cohesion was intentionally manipulated by adding or deleting connectives and overlapped lexical expressions in the texts (e.g., McNamara et al., 2011). There were two versions of texts: high-cohesion and low-cohesion texts, and after reading the text, the participants were asked to rate how cohesive the text was, to measure whether the readers could judge the degree of cohesion of the text appropriately.

4.1.2 Method

4.1.2.1 Participants

A total of 30 Japanese EFL undergraduate and graduate students participated in this experiment (21 males and nine females; average age = 22.59, range = 19–27). The undergraduates were majoring in various fields, including agro-biological resource science, art and design, engineering science, humanities, international studies, mathematics, science and engineering. The graduate participants were majoring in comprehensive human sciences, education, humanities and social sciences, life and environmental sciences, pure and applied sciences, and systems and information engineering. None of them had physical disabilities that interfered with the activity of

reading.

The native language of the participants was Japanese, and they had studied English for at least six years in the Japanese education system. Their general English proficiency level was estimated to be from beginner to proficient (i.e., CEFR A1 to C1 level) based on their self-reported standardized test scores on the following tests: the TOEIC listening and reading test (M = 750.45, SD = 152.38, range 520 to 970, n = 11), the TOEFL ITP test (M = 503.00, SD = 2.45, range 500 to 506, n = 3), the IELTS (M = 6.00, n = 2), and the EIKEN test (Grade 4 to Pre-1: Grade 4, n = 2; Grade 3, n = 2; Grade Pre-2, n = 2; Grade 2, n = 5; Grade Pre-1, n = 6). Note that 25 participants reported at least one of the scores, and five participants reported none.

The participants were classified into two reading proficiency groups (high and low proficiency) based on their test scores on an L2 reading proficiency test conducted in the experiment. Proficiency was considered a factor in the analysis of the results. Data from all participants were analyzed.

4.1.2.2 Materials

Experimental texts

Two expository texts about scientific phenomena were used in this experimental study (see Appendix A–D). For the manipulation of cohesiveness, two texts (*Heat Changes Matter* [Heat], *Space Travel and the Human Body [Space]*) were adopted from previous an EFL reading study (Hosoda, 2016), which originally came from other prior studies (e.g., McNamara et al., 2011).

The Heat text explains the three states of matter (i.e., solid, liquid, and gas); The Space text explains how the effects of gravity on the human body change in space travel. These passages were chosen because: (1) both explain a scientific phenomenon, (2) their

content and difficulty are appropriate for the target readers, and (3) the texts were used in studies investigating the effect of text cohesion on Japanese EFL readers.

The texts were revised into two cohesion versions (high-cohesion and low-cohesion) based on the criteria of McNamara et al. (2011) and Ozuru et al. (2009). The following revisions were conducted in the high-cohesion texts: (a) explication of pronouns, (b) explication of omitted information, and (c) addition of discourse markers. An example of a high-cohesion text is provided in Table 4.1.

Table 4.1

Example of Experimental the Heat Text in Experiment 1

Adding heat or taking away heat can change matter. Matter can change **from one state to another state, or from one form to another form**. Three states of matter are solid, liquid and gas. **For example**, an ice cube is solid water. **Heat can melt an ice cube**, causing the ice cube to change into liquid water. When heat is taken away, the liquid water can change back into solid water. Heat can make liquids boil **and change into a gas state**. **For example**, water boils when it is heated. As the water boils, it turns into a gas state that is called water vapor. Heat from the sun causes liquid water to turn into water vapor. Water vapor **then** mixes with the air in a process called evaporation. **However**, sometimes heat causes irreversible changes. **As one example**, bread can change into toast when you heat **the bread**. **However**, you cannot untoast a piece of toast **by taking away heat**. **As another example**, eggs change when you cook them **in a pan**. **However, of course** you cannot uncook an egg **by taking away the heat**.

Note. Bold letters were added or revised only in the high-cohesion text. In the low-cohesion text, the bold parts were rephrased to simpler expressions, or deleted.

This revision was also conducted in prior studies (Hosoda, 2016), and differences between high- and low-cohesion texts were confirmed. Materials from these studies were used in this study.

All texts were checked for word difficulty in the EFL reading. Low-frequency words of a level of 5 or higher in the *Japan Association of College English Teachers (JACET) 8000 List* (JACET 8000 Basic Word Revision Committee, 2003) were replaced with higher-frequency synonyms or annotated in Japanese to make the text readable for EFL readers. This was conducted to prevent the experimental texts from having different levels of difficulty in terms of vocabulary for each cohesion version. In addition, a native English speaker checked the materials for their native acceptability. Table 4.2 presents the textual features of the experimental texts.

Table 4.2

Textual Features of Experimental Texts Used in Experiment 1

	Heat		Space	
Text Feature	Low-cohesion	High-cohesion	Low-cohesion	High-cohesion
Words	157	185	128	177
Sentences	22	16	11	11
FKGL	3.7	5.8	5.3	7.6
FRE	81.0	73.9	77.3	68.6

Note. FKGL = Flesch-Kincaid grade level; FRE = Flesch Reading Ease.

Cohesion-judgment task

To measure participants' awareness of the degree of cohesion, a cohesion-judgment task was created based on previous studies (Helder et al., 2016; Nahatame, 2017). This task asked participants whether the text was cohesive or not after reading it, through the question, "*Was the text cohesive by containing explicit information?*" Readers were asked to rate this cohesion through a six-point Likert scale (1: *I never think so* to 6: *I think so very much*). Appendix E presents the actual directions in Japanese.

Participants were asked to carefully read the text for understanding with no time limit. They read two experimental texts, one in the high-cohesion condition, and the other in the low-cohesion version. The text assignments were counterbalanced randomly. Participants read freely without a time limit, but all of them completed the reading of their text within 10 minutes. Participants completed the cohesion-judgment task after reading each text and graded each text's cohesion via a six-point Likert scale.

Reading strategy questionnaire

Additionally, to investigate participants' use of reading strategies in ordinary reading as supplementary data, a 30-item questionnaire was adopted from the metacognitive awareness of reading strategies inventory (MARSI) of Mokhtari and Sheorey (2002). In this questionnaire, there were three types of strategies (i.e., global, support, and problem-solving). The details of each category are as follows: First, the global strategies category contains 13 items and represents a set of reading strategies oriented toward a global analysis of the text (e.g., "I critically analyze and evaluate the information presented in the text"). Second, the support strategies category contains nine items and involves the use of outside reference materials, taking notes, and other practical strategies (e.g., "I go back and forth in the text to find relationships among ideas in it").

Third and lastly, the problem-solving strategies category contains eight items related to solving problems when a text becomes difficult to read (e.g., “I try to guess the meaning of unknown words of phrases”). All items included in the questionnaire are listed in Table 4.3. Participants answered each item, written in Japanese, on a five-point Likert scale (1: *I never or almost never do this*. to 5: *I always or almost always do this.*). Appendix F shows the actual questionnaire written in Japanese that was presented to the participants.

Table 4.3

Reading Strategy Questionnaire Adopted from Mokhtari and Sheorey (2002) in Experiment 1

Item	Type	Strategy
1	GLOB	I have a purpose in mind when I read.
2	SUP	I take notes while reading to help me understand what I read.
3	GLOB	I think about what I know to help me understand what I read.
4	GLOB	I preview the text to see what it's about before reading it.
5	SUP	When text becomes difficult, I read aloud to help me understand what I read.
6	SUP	I summarize what I read to reflect on important information in the text.
7	GLOB	I think about whether the content of the text fits my reading purpose.
8	PROB	I think about whether the content of the text fits my reading purpose.
9	SUP	I discuss what I read with others to check my understanding.
10	GLOB	I skim the text first by noting characteristics like length and organization.
11	PROB	I try to get back on track when I lose concentration.
12	SUP	I underline or circle information in the text to help me remember it.
13	PROB	I adjust my reading speed according to what I'm reading.
14	GLOB	I decide what to read closely and what to ignore.
15	SUP	I use reference materials such as dictionaries to help me understand what I read.
16	PROB	When text becomes difficult, I pay closer attention to what I'm reading.
17	GLOB	I use tables, figures, and pictures in text to increase my understanding.
18	PROB	I stop from time to time and think about what I'm reading.
19	GLOB	I use context clues to help me better understand what I'm reading.
20	SUP	I paraphrase (restate ideas in my own words) to better understand what I read.
21	PROB	I try to picture or visualize information to help remember what I read.
22	GLOB	I use typographical aids like boldface and italics to identify key information.
23	GLOB	I critically analyze and evaluate the information presented in the text.
24	SUP	I go back and forth in the text to find relationships among ideas in it.
25	GLOB	I check my understanding when I come across conflicting information.
26	GLOB	I try to guess what the material is about when I read.
27	PROB	When text becomes difficult, I reread to increase my understanding.
28	SUP	I ask myself questions I like to have answered in the text.
29	GLOB	I check to see if my guesses about the text are right or wrong.
30	PROB	I try to guess the meaning of unknown words or phrases.

Note. GLOB = global; SUP = support; PROB = problem-solving. All items were translated into Japanese in the experimental study.

L2 reading proficiency test

This study used participants' English reading proficiency level as a factor in the analysis. To examine participants' proficiency, a 24-item L2 reading proficiency test was adopted from the reading sections of the EIKEN test (Obunsha, 2003b, 2009). This test requires readers to understand the content of passages and not to demonstrate grammatical or lexical knowledge. It had a multiple-choice format and was conducted for 30 minutes.

Six items were extracted from the pre-first-grade test, and 18 items from the second-grade test. The EIKEN test was used to measure participants' proficiency because it has been widely used in experiments assessing Japanese EFL learners' English language reading proficiency (e.g., Ushiro et al., 2018).

Participants were told to read the test passages in accordance with the order presented in the booklet, and answer the multiple-choice questions.

Moreover, they were told to answer either *Yes* or *No* to the question, "Have you read this test passage before?" upon reading it. This was to confirm that they had not read the passage before the experiment.

4.1.2.3 Procedure

The participants were tested individually in four sessions lasting approximately 50 minutes, a pre-study session, test session, and reading session. The procedure for the second study is shown in Figure 4.1.

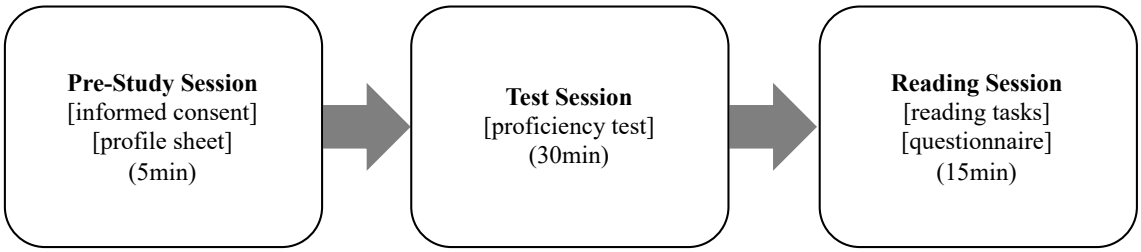
In the pre-study session, before starting the experiment, the researcher explained the study's purpose, procedure, and method of data collection, and informed consent was obtained. If participants agreed to participate in the experiment, they were asked to fill in a demographic profile sheet. The items on the profile sheet were name, age, sex, department and year of university, English qualifications (e.g., TOEIC score), and

experience studying abroad. These questions were used to confirm the participants' English study experience.

Subsequently, the experiment progressed to the test session phase. In the test phase, participants took the L2 reading proficiency test with a time limit of 30 minutes. After the proficiency test, a five-minute rest time was given. Then, in the reading session, they read two passages in different cohesion conditions, and completed the cohesion-judgment task. Finally, the participants completed the 30-item reading strategy questionnaire in five minutes.

Figure 4.1

Procedure of Experiment 1



Note. proficiency test = L2 reading proficiency test; reading tasks = reading texts and cohesion-judgment task; questionnaire = reading strategy questionnaire.

4.1.2.4 Scoring and Data Analysis

Cohesion-judgment task

Before analysis, the researcher summarized the scores ranging from 1 to 6 for the cohesion-judgment tasks. Then, the means of the judgment scores for cohesion were calculated. A two-way ANOVA was conducted on the judgment rates, with participants' reading proficiency as a between-participant variable and cohesion condition as a within-participant variable.

Reading strategy questionnaire

To obtain the general tendencies and reading profiles of the participants, the scores of the reading strategy questionnaire were summarized in terms of three categories of strategies (global, support, and problem-solving) and the readers' proficiency. Descriptive statistics of the results were confirmed.

L2 reading proficiency test

The L2 reading proficiency test contained 24 multiple-choice questions. If participants answered one item correctly, they were given one point. Possible scores ranged from 0 to 24.

4.1.3 Results

4.1.3.1 L2 Reading Proficiency Test

The reliability of the L2 reading proficiency test was sufficient (Cronbach's $\alpha = .74$) after excluding one item because of its low discriminability. Before the main analysis, participants were classified into two reading proficiency groups: high-proficiency and low-proficiency groups. Table 4.4 shows the descriptive statistics of the L2 proficiency test results.

As shown in Table 4.4, the median test score (i.e., 14/23) divided participants into a high-proficiency group ($n = 19$) and a low-proficiency group ($n = 11$). The mean score of the high-proficiency group ($M = 15.55$, $SD = 1.79$) was significantly higher than that of the low-proficiency group ($M = 10.09$, $SD = 2.84$), $t(28) = 6.72$, $p = .017$, $r = .79$, with a large effect size.

Table 4.4

Means With 95% CIs and SDs of the Proficiency Test Results in Experiment 1

Proficiency	*N*	*M*	95%CI	*SD*	*Min*	*Max*
High	19	15.55	[14.74, 16.36]	1.79	14.00	20.00
Low	11	10.09	[8.41, 11.77]	2.84	5.00	13.00
Total	30	13.63	[12.38, 14.88]	3.49	5.00	20.00

Note. CI = confidence interval; *SD* = standard deviation. Maximum possible score was 23.00.

4.1.3.2 Cohesion-Judgment Task

Table 4.5 and Figure 4.2 show the descriptive statistics of the results of the cohesion-judgment task, which is the main focus of this experiment. Generally, participants rated a higher score after reading the high-cohesion text than the low-cohesion text (high-cohesion: $M = 4.77$, $SD = 1.10$; low-cohesion: $M = 4.57$, $SD = 0.97$). In addition, the rating values tended to be higher among high-proficiency readers than among low-proficiency readers (high-cohesion: $M = 4.84$, $SD = 0.90$; low-cohesion: $M = 4.68$, $SD = 0.95$ among high-proficiency readers; high-cohesion: $M = 4.64$, $SD = 1.43$; low-cohesion: $M = 4.36$, $SD = 1.03$ among low-proficiency readers).

Table 4.5

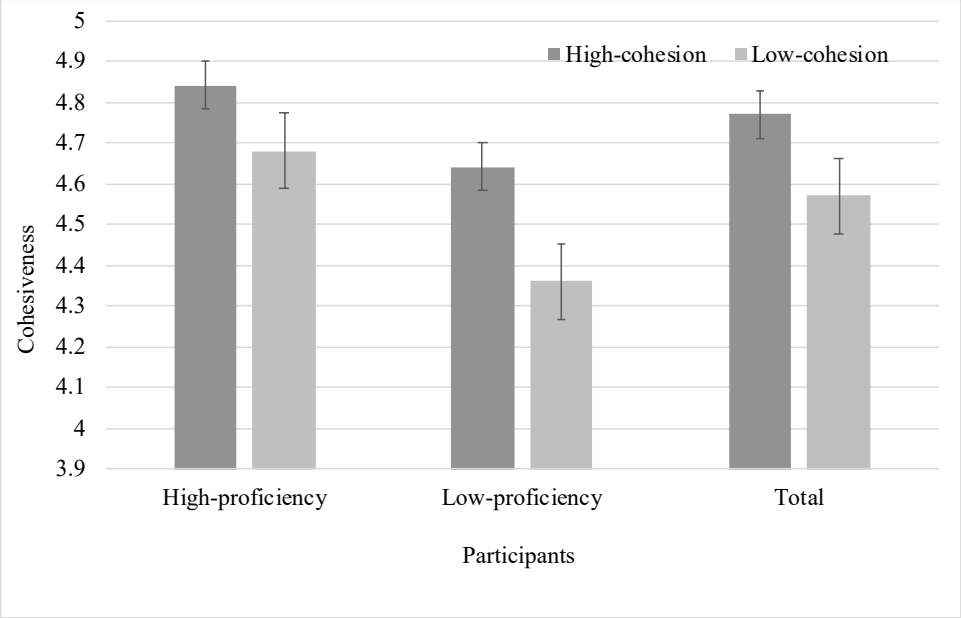
Means With 95% CIs and SDs of the Cohesion-Judgment Task Results in Experiment 1

		High-Cohesion			Low-Cohesion		
Proficiency	*n*	*M*	95%CI	*SD*	*M*	95%CI	*SD*
High	19	4.84	[4.44, 5.25]	0.90	4.68	[4.26, 5.11]	0.95
Low	11	4.64	[3.79, 5.48]	1.43	4.36	[3.76, 4.97]	1.03
Total	30	4.77	[4.37, 5.16]	1.10	4.57	[4.22, 4.91]	0.97

Note. Maximum possible score was 5.00.

Figure 4.2

Means of Cohesion Rating on High- and Low-Cohesion Texts by Readers' Proficiency in Experiment 1



Note. Error bars represent the standard errors of the means.

A 2 (cohesion: high-cohesion, low-cohesion) × 2 (proficiency: high-proficiency, low-proficiency) two-way mixed ANOVA was conducted on the mean rate of the

cohesion-judgment task to confirm whether there was a significant difference between high-proficiency readers and low-proficiency readers in terms of text cohesion judgment, with cohesion as a within-participant variable and the proficiency as a between-participant variable. The results of ANOVA are summarized in Table 4.6. The results showed that there was no significant main effect for cohesion, $F(1, 28) = 0.71$, $p = .405$, $\eta^2 = .01$, or proficiency, $F(1, 28) = 0.74$, $p = .400$, $\eta^2 = .02$. Additionally, the Cohesion × Proficiency interaction was not significant, $F(1, 28) = 0.05$, $p = .823$, $\eta^2 < .01$. Note that all effects had an extremely small effect size.

These results indicate that the participants were not able to clearly judge the different levels of text cohesion of the low- and high-cohesion texts, suggesting that Japanese EFL readers were not sensitive to the degree of cohesiveness of a text (i.e., they did not monitor text cohesion).

Table 4.6

ANOVA Results for Main Effects and Interaction Effects of Cohesion and Proficiency on the Mean Rate of Cohesion-Judgment Task in Experiment 1

Source	*SS*	*df*	*MS*	*F*	*p*	η^2
	Between participants					
Proficiency (P)	0.97	1.00	0.97	0.73	.396	.02
Error	36.37	28.00	1.30			
	Within participants					
Cohesion (C)	0.65	1.00	0.65	0.71	.405	.01
C × P	0.05	1.00	0.05	0.05	.823	<.01
Error (C)	25.35	28.00	0.91			
Total	63.38	29.00				

4.1.3.3 Reading Strategy Questionnaire

The reliability of the reading strategy questionnaire was sufficient (Cronbach's α = .70). Table 4.7 presents the descriptive statistics of the questionnaire results. Note that participants answered each item on a five-point Likert scale (1: *I never or almost never do this*. to 5: *I always or almost always do this.*).

To check the tendency of reading strategy use in each proficiency group, the mean rates of 30-item strategies (see Table 4.7 or Appendix F) were compiled for the three categories of strategy (i.e., global, problem-solving, and support reading strategies) and reader proficiency. As the rated numbers (e.g., 3: *I sometimes do*, 4: *I usually do*) in Table 4.7 show, readers of both groups used reading strategies relatively frequently when they read text written in English. This suggests that the participants in the different proficiency groups were not very different in recognizing their use of strategies.

Table 4.7

Means With 95% CIs and SDs of the Reading Strategy Questionnaire Results in Experiment 1

Proficiency	C	*n*	*M*	95%CI	*SD*	*Min*	*Max*
High	G	19	3.49	[3.22, 3.76]	0.56	2.31	4.31
	P	19	3.57	[3.45, 3.68]	0.24	3.12	4.00
	S	19	3.43	[3.15, 3.70]	0.56	2.33	4.56
Low	G	11	3.52	[3.34, 3.69]	0.26	3.00	3.85
	P	11	3.69	[3.41, 3.98]	0.42	2.75	4.12
	S	11	3.42	[3.08, 3.77]	0.52	2.44	4.22
Total	G	30	3.50	[3.33, 3.68]	0.47	2.31	4.31
	P	30	3.61	[3.49, 3.73]	0.32	2.75	4.12
	S	30	3.43	[3.22, 3.63]	0.54	2.33	4.56

Note. C= category; G = global reading strategies; P = problem-solving strategies: S = support reading strategies. Maximum possible score was 5.00.

4.1.4 Discussion

Response to RQ2-1: *Are EFL readers aware of the degree of cohesiveness of a text?*

The result of a two-way ANOVA of the cohesion-judgment task for two cohesion conditions showed that there were no significant effects of text cohesion and participants' proficiency in rating text cohesion of the text. This finding suggested that participants could not judge the texts' degree of cohesion.

However, taking a look at the descriptive statistics of the cohesion-judgment task (Table 4.5), participants rated a higher score after reading the high-cohesion text than the low-cohesion text. Although the participants' judgment was not significantly different

between cohesion conditions (i.e., high-cohesion and low-cohesion), the participants were possibly able to make a different grade depending on the degree of cohesion when they read the various cohesive texts. Since the present study explored readers' awareness regarding text cohesion exploratory, it was difficult to compare our results with those of other empirical studies. However, our findings could be interpreted in light of the following two possible explanations:

First, the participants in this study might have focused mainly on the meaning evoked in the text, and they might not have focused on text cohesion as a textual (linguistic) aspect. Indeed, they were told to read the experimental texts for their understanding, not being instructed to pay attention to text cohesion. Thus, it is possible that the readers' standards of coherence (e.g., Graesser et al., 1994; van den Broek et al., 2001) were not directed to meta-linguistic features.

In previous studies that conducted experimental studies to explore the effect of text cohesion on reading, it was revealed that text cohesion affects readers' processing and understanding of the text (e.g., McNamara, 2001; McNamara et al., 1996; Horiba, 1996; Hosoda, 2016; O'Reilly & McNamara, 2007). This is because text cohesion is a critical factor for understanding various aspects of a text: understanding the literal meaning of the text, understanding the relationship between ideas in the text, and making inferences from the text. While text cohesion affects comprehension, it may not always be clearly perceived by the reader. It is also important to note that cohesion is a continuum, not a binary value that can be clearly distinguished for everyone, such as high or low-cohesion. It is possible that the experimental environment, in which no specific instructions were given, affected the judgments of the participants on cohesion and that they did not clearly differentiate between high and low-cohesion.

Second, it might be inappropriate to use a text cohesion-judgment task to measure

readers' awareness regarding texts' degree of cohesiveness. This task was constructed by the researcher to directly compare text cohesion and coherence. Although this judgment task has been used in several studies (e.g., Helder et al., 2016; Nahatame, 2017; Tadoro et al., 2010), the focus of these studies was coherence, or whether the text content contains inconsistent information, and whether the text makes sense. Hence, the reliability and validity of the cohesion-judgment task in the present study might not be sufficient. In other words, this task was not adequate to reflect readers' judgments of cohesion. Therefore, there is still room to investigate readers' cohesion precisely because the participants in this study were actually graded at different rates in the task.

Response to RQ2-2: ***Do EFL readers judge high- and low-cohesion texts differently in terms of their proficiency level?***

Based on the ANOVA results discussed above, it is suggested that readers' proficiency does not affect their ratings in the cohesion-judgment task. In particular, neither analysis displayed a significant simple effect of proficiency or interactional effects for Text Cohesion × Proficiency. However, previous studies have pointed out the effect of readers' age and proficiency on monitoring skills for text coherence or their understanding (e.g., Nahatame, 2017; Takaki, 2014; Todaro et al., 2010). Note that these studies has not directly addressed text cohesion, so caution is needed in interpretation. This inconsistency between the current study and previous studies could be attributed to two possible reasons.

First, it could be true that there was a lack of variance in readers' proficiency, and thus the effect of proficiency was difficult to observe. The participants in this study were all graduate or undergraduate students at the same university, and they had the necessary level of skill to pass the university entrance exam. The participants were proficient in

English and not so different in terms of their reading proficiency. This may be why there were no significant effects on their proficiency in performance in the cohesion and coherence judgment tasks.

Second, participants were instructed as follows: "Carefully read the text for your understanding" in this experiment. This instruction may have facilitated readers' employment of high standards of coherence. Standards of coherence vary depending on readers' purpose (e.g., Graesser et al., 1994; Linderholm et al., 2002; van den Broek et al., 1995; van den Broek et al., 2001). Thus, by receiving this instruction, readers might have tried to carefully capture the situation described in the text, causing no statistical differences between the two proficiency groups to be observed. Since this study did not directly measure readers' comprehension, future studies need to address this matter.

4.1.5 Conclusion of Experiment 1

The purpose of Experiment 1 was to investigate Japanese EFL readers' judgments of text cohesion in reading and two RQs were set: *Are EFL readers aware of the degree of cohesiveness of a text?* (RQ2-1) and *Do EFL readers judge high- and low-cohesion texts differently in terms of their proficiency level?* (RQ2-2). To discover readers' judgment of text cohesion and its relation to their English proficiency, one experiment was conducted on 30 Japanese university students using a cohesion-judgment task.

The results revealed that Japanese EFL college-level readers could discern differences in text cohesiveness, but the cohesion-judgment rates between high- and low-cohesion texts were not significantly different (RQ2-1). In addition, no effect of readers' proficiency on cohesion-judgment was observed (RQ2-2).

Given the results of Experiment 1, Japanese EFL readers could not detect the degree of text cohesiveness at a statistically significant level. Although text cohesion is a critical

factor in reading (Halliday & Hasan, 1976), it seemed to be difficult for these Japanese EFL readers to monitor while reading. However, even if the reader is unaware of text cohesiveness, the latter in English texts can affect EFL reading (e.g., Hosoda, 2016). Since Experiment 1 did not assess readers' comprehension, whether and how the text cohesiveness affects EFL reading should be explored. Therefore, the following experiments investigated whether the different cohesive texts affected Japanese EFL learners' comprehension of expository texts.

4.2 Experiment 2: Effects of Text Cohesion on Japanese EFL Readers' Comprehension of Expository Texts on Social Topics

4.2.1. Experiment 2-A

4.2.1.1 Purpose and Research Questions

This experimental study investigated the effects of text revision on elaborate text content by using the same words or adding conjunctions (e.g., Vidal-Abarca et al., 2000), and comparing high-cohesion text to low-cohesion text. According to previous studies, it is assumed that low-proficiency readers who have a problem in inferring ideas within the text show good comprehension in high-cohesion text, whereas high-proficiency readers show better comprehension in the low-cohesion text. However, some studies have reported that high-proficiency readers also comprehend high-cohesion texts better (e.g., Voss & Silfies, 1996). Thus, it is arguable whether the revised high-cohesion text will improve different proficiency readers' comprehension. Moreover, previous studies have not sufficiently investigated Japanese EFL learners. There is room to investigate text revision for EFL learners who are thought to have more problems in reading than native speakers. Thus, Experimental Study 2 aimed to examine the effect of text revision on Japanese EFL learners and uncover which type of text (high-cohesion/low-cohesion) facilitated readers' comprehension. Experiment 2 consisted of two experiments (Experiment 2-A and 2-B). Experiment 2-A was conducted with relatively high-proficiency learners, and Experiment 2-B was a follow-up experiment, conducted with relatively low-proficiency learners. The following research questions were investigated in experimental study 2.

RQ3-1: Does high or low cohesion affect Japanese EFL readers' surface-level comprehension of social expository texts?

RQ3-2: Does high or low cohesion affect Japanese EFL readers' deep-level comprehension of social expository texts?

RQ3-3: Does the cohesion effect on the comprehension of social expository texts vary as per the readers' proficiency level?

University students were chosen as participants to reveal the effect of text revision in terms of cohesion for Japanese EFL learners. These participants were selected for practical reasons for the experimental study. Moreover, prior studies of EFL readers (Hosoda, 2016) have used participants of similar age, which makes generalizations across studies possible.

Participants were asked to take a written recall task as a measure of textbase-level comprehension and why-questions as an assessment of situation-model-level comprehension. Additionally, learners' impressions of the texts and reading strategies were evaluated using questionnaires as supplemental data. In addition, to measure and consider the results in the context of learners' reading ability, an L2 reading proficiency test was conducted.

4.2.1.2 Method

4.2.1.2.1 Participants

Twenty-four Japanese EFL undergraduate students participated in Experiment 2-A (16 males and eight females; average age = 20.32, range = 19–22). Their native language was Japanese, and they had studied English for at least six years in the Japanese education system. Their university majors were varied, including agro-biological resource science, chemistry, education, engineering science, health and physical education, humanities, information science, Japanese language and culture, mathematics, medicine, nursing, and

social studies. They were classified into two reading proficiency groups (high-proficiency and low-proficiency groups) based on their L2 reading proficiency test scores.

4.2.1.2.2 Materials

Experimental texts

There were four experimental texts: two original expository texts and two cohesion versions (high-cohesion, low-cohesion texts) of each used in this experimental study (see Appendix G–J). The two original texts ("Two-for-one; TFO," "Fast Food for Bears; FFB") were adopted from the reading section of the pre-second grades of the EIKEN test (Obunsha, 2003a, 2006). First, based on Vidal-Abarca et al. (2000), two versions of the original texts (high-cohesion and low-cohesion texts) were created in terms of text cohesion. Vidal-Abarca et al.'s text revision was used here because they examined cohesion manipulation for social- or historical-topic texts. The following revisions were made to the revised high-cohesion version: (a) explication of pronouns, (b) explication of omitted information, and (c) addition of discourse markers.

An example of manipulation is as follows: in the low-cohesion text, *In Peru all children need to attend primary school. In reality 23 percent of students stop going to school before the fifth grade. They find the classes too difficult.*, and in the high-cohesion text, *In Peru all children need to attend primary school. However, in reality 23 percent of students stop going to school before the fifth grade. This is often because the children find the classes too difficult.* As another instance, the low-cohesion text describes *Sometimes their parents themselves cannot read and write. This means that the children have little chance at home to learn the basic skills necessary to understand their classes*, and the high-cohesion text conveys, *Sometimes their parents themselves cannot read and write. This means that the children have little chance at home to learn the basic skills of reading*

and writing necessary to understand their classes.

Second, to prevent EFL learners from having difficulty reading, all texts were checked in terms of word difficulty. To make the text readable, low-frequency words of level 5 or higher in *JACET 8000 List* (JACET Basic Word Revision Committee, 2003) were replaced with higher-frequency synonyms or annotated. Finally, a native English speaker checked the materials for their native acceptability. Table 4.8 shows the textual features of each text version of both texts.

Table 4.8

Textual Features of Experimental Texts Used in Experiment 2

	Low-Cohesion			High-Cohesion		
Text	Words	Sentences	FKGL	Words	Sentences	FKGL
TFO text	273	17	8.4	321	17	9.6
FFB text	297	15	8.1	343	16	9.0

Note. TFO = "*Two-for-one*," FFB = "*Fast Food for Bears*," FKGL= Flesch-Kincaid grade level.

In the experiment, all participants were asked to read two texts in different conditions (i.e., high-cohesion/low-cohesion). Before reading the first text, participants were informed that they would have to recall the text content and answer some comprehension questions without looking at it again. They were asked to read the text carefully with no time limit.

The assignments of both texts were randomly counterbalanced. Participants were provided with one of the following four text assignments: (a) provided revised TFO text first and control FFB text second, (b) provided control TFO text first and revised FFB

text second, (c) provided revised FFB text first and control TFO text second, and (d) provided control FFB text first and revised TFO text second.

Written recall task

A free written recall task was conducted to determine participants' textbase-level comprehension, (see Appendix K). Written recall tasks have been used to assess reading outcomes resulting from interactions between readers and text (e.g., Koda, 2005). This task measures readers' memory of explicit ideas stated in the text and requires readers to write down as much information about the text as they can. After reading, the participants reproduced the text content without referring to the text.

The participants completed the task in Japanese, because previous studies have suggested that it is not valid to use production of L2 (i.e., English in this case) to assess the understanding of L2 (e.g., Davis & Bistodeau, 1993). Even if readers comprehend the text content, it is possible that they are not able to demonstrate their understanding of L2 because of a lack of L2 writing proficiency. Therefore, the written recall task was conducted in Japanese. In addition, it was performed under no time limit to allow readers to write all the information they remembered.

Why-question task

To measure participants' situation-model-level comprehension, a why-question task was conducted. One question was asked for each text. This task was used because the written recall task was not able to assess readers' deeper understanding of the text (McNamara et al., 1996). The why-question task was performed to determine readers' situational understanding and generation of bridging inferences. This task required readers to describe the causal event in the text (e.g., "Why *A Fed Bear is a Dead Bear*

Campaign influenced positively on both human and bears?"). If readers understand the text coherently, they can answer this question easily based on the inferences they generate during reading. In this study, the participants were asked to explain the causal relationships communicated through the text. Appendix L presents the materials used for this task.

In addition to the written recall task, participants were asked to answer the why-question in Japanese. There was no time limit for this task, but they had to complete it without referring to the experimental texts.

Questionnaire

As supplemental data, eight question items were used to check the readers' impressions of each text (e.g., difficulty, coherence of the text) and learners' reading strategies used during reading (see Appendix M). The specific items were as follows: (1) "Were you able to read the text precisely?" (2) "Was the text easy to read?" (3) "Were you able to visualize events and scenes written in the text while you were reading?" (4) "Was the text clear and obvious?" (5) "Did you have problems in comprehending the text?" (6) "Did you especially pay attention to detail information?" (7) "Were you able to recall the text content smoothly and answer the questions?" and (8) "Was the information explained in the text coherent and easy-to-grasp?." Questions (1), (5), and (7) measured readers' achievement of the reading and the task. Questions (2), (4), and (8) measured the readability of the text. Questions (3) and (6) measured readers' strategy use during reading.

A five-point Likert scale was used to record participants' impressions using the following words: 5: "*I agree*," 4: "*I tend to agree*," 3: "*I am neutral*," 2: "*I tend to disagree*," 1: "*I disagree*." Participants rated the questionnaire by circling their preferred response (1–5) after reading each experimental text. There was no time limit.

L2 reading proficiency test

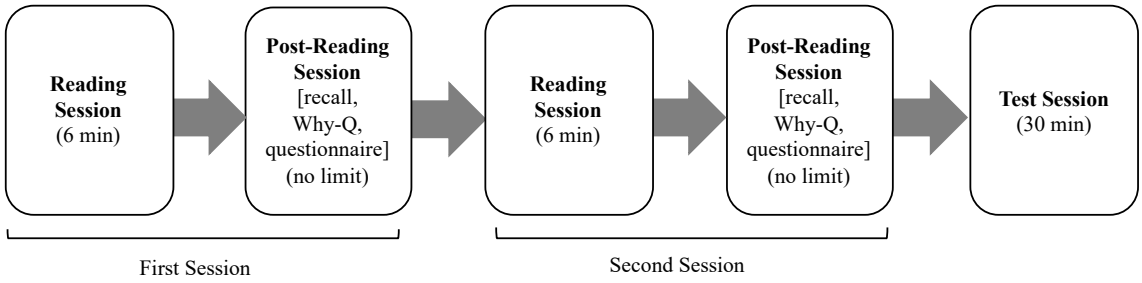
This study used participants' English reading proficiency level as one factor of analysis. To examine participants' L2 reading proficiency, a 24-item test was adopted from the reading section of the pre-first and second grades of the EIKEN test (Obunsha, 2003b, 2009). As explained in previous sections, the EIKEN test was used because it has been widely adopted in experiments to assess Japanese EFL learners' reading proficiency in English. Six items were extracted from the pre-first grade test, and 14 items were extracted from the second grade of the test. The grade selected was based on the expected English proficiency of the participants.

4.2.1.2.3 Procedure

Experiment 2-A was conducted in a session of about 100 minutes and included four phases: a pre-study session, reading session (first and second), post-reading session (first and second), and test session (see Figure 4.3).

Figure 4.3

Procedure of Experiment 2-A



Note. recall = written recall task; why-Q = why-question task; proficiency test = L2 reading proficiency test.

In the pre-study session, the researcher explained the purpose and procedures of the experimental study. If participants agreed to participate in the experiment, they were asked to fill in a profile sheet for demographic information. Items in the profile sheet were: name, age, sex, department and grade of university, qualification of English (e.g., TOEIC score, EIKEN test grade, TOEFL score), and experience of studying abroad. These questions were used to confirm the participants' experience of studying English. After receiving informed consent from the participants, the experiment progressed to the reading session phase.

In the reading session, participants read one experimental text either in the high or low-cohesion condition, within six minutes. After reading, they completed the written recall task within 20 minutes, why-question task within 13 minutes, and the questionnaire of the post-reading session. Since there were two experimental texts, this procedure was repeated (i.e., first and second session in Figure 4.3). Finally, in the test session, participants took the L2 reading proficiency test, with a time limit of 30 minutes.

4.2.1.2.4 Scoring and Data Analysis

L2 Reading Proficiency Test

The L2 reading proficiency test contained 24 multiple-choice questions. If participants' answers were correct, they were given 1 point. Possible scores ranged from zero to 24.

Written Recall Task

Before scoring, the experimental texts were divided into idea units (IUs) based on the criteria of Ikeno (1996). Ikeno's criteria were used because other L2 reading studies

employed these criteria (e.g., Ushiro et al., 2007). The IU division was arranged in the following criteria: (a) IU consisted of a single clause, (b) infinitive, particle construction, gerundive, nominalized verb phrase, and heavy adjunct were considered as separate IUs, (c) conjuncts and disjuncts were separated into different IUs.

This IU-division was conducted by two undergraduate students who studied English education at university. They divided the texts into IUs individually. The inter-rater agreement for IU-division was 91.09%. Disagreements between the two raters were resolved through discussion. The number of IUs was 67 for FFB (control) text, 80 for FFB (revised) text, 61 for TFO (control) text, and 77 for TFO (revised) text.

In the recall data scoring, 30 percent of the recall data was scored by the two raters. They scored the data individually. If two-thirds of the IU information was recalled, one point was given. Inter-rater agreement was 80.62%, and all disagreements were resolved through discussion. After the scoring criteria were finalized, the remaining 70 percent of the data were scored by the researcher alone. Finally, the recall production rate was calculated. In addition, to compare the recall production according to each text, an arcsine transformation was performed on the total recall production rates because each text had a different number of IUs. To examine the effect of text revisions on textbase-level comprehension, the final recall production was analyzed using a two-way mixed ANOVA, with participants' reading proficiency as a between-participant variable and cohesion condition as a within-participant variable.

Why-question Task

In this experimental study, why questions were scored based on the existence of the key or important ideas of the text used to answer the causal questions (e.g., "Why *A Fed Bear is a Dead Bear Campaign* influenced positively on both human and bears?"). Before

the scoring, two raters decided on five important ideas that were expected to be answered to explain the causal questions. Both raters discussed whether the five ideas expressed cause-effect relations to answer the why-questions, creating model answer descriptions based on these ideas. Then, one key idea that had the core information for the answer was chosen (e.g., “Campaigners say that the best way to protect the bears is to stop bears from eating garbage by using garbage containers that are too difficult for bears to open.”). If the key idea was in the participants’ answer, one point was given. In addition to the key idea, if participants wrote other ideas (e.g., “When people do this, the bears leave, and this makes life safer for both animals and human beings.”), they were given one point for each idea.

Two raters studying English education in universities scored 30 percent of the why questions separately. The inter-rater agreement was 85.00%, and disagreements were resolved through discussion. The remaining data were scored by the researcher alone. Finally, the sums of the points, ranging from zero to five, were calculated.

To allow statistical comparison, participants’ answers to the why-question were further divided into three groups depending on their coherence (i.e., low, medium, and high coherent understanding group). The low group contained the participants who received zero or one point, with the medium group containing those who received two or three points, and the high group those who received four or five points. The number of people in the three groups was compared by text type (high-cohesion/low-cohesion). To investigate the relationship between readers’ comprehension, Fisher’s exact test was conducted.

4.2.1.3 Results

4.2.1.3.1 L2 Reading Proficiency Test

The reliability of the L2 reading proficiency test was sufficient (Cronbach's α = .81) after excluding three items because of their low discriminability. Before the main analysis, the participants were classified into two reading proficiency groups. Table 4.9 shows the descriptive statistics of the L2 proficiency test results. As shown in Table 4.9, the median test score divided participants into a high-proficiency group (n =13) and a low-proficiency group (n = 11). The mean score of the high-proficiency group (M = 16.92, SD = 1.85) was significantly higher than that of the low-proficiency group (M = 10.00, SD = 2.76), $t(22) = 7.33$, $p < .001$, $d = 3.00$.

Table 4.9

Means With 95% CIs and SDs of the Proficiency Test Results in Experiment 2-A

Proficiency	*n*	*M*	95%CI	*SD*	*Min*	*Max*
High	13	16.92	[15.81, 18.04]	1.85	14.00	20.00
Low	11	10.00	[8.15, 11.85]	2.76	5.00	13.00
Total	24	13.75	[11.98, 15.52]	4.18	5.00	20.00

Note. CI = confidence interval; SD = standard deviation. Maximum possible score was 21 (24 items minus three items with low discriminability).

4.2.1.3.2 Written Recall Task

Table 4.10 and Figure 4.4 show the mean recall rate and standard deviations for each text in the written recall task. Before the analysis, arcsine transformation was performed on the total recall production rates because number of IUs differed in each text.

Table 4.10

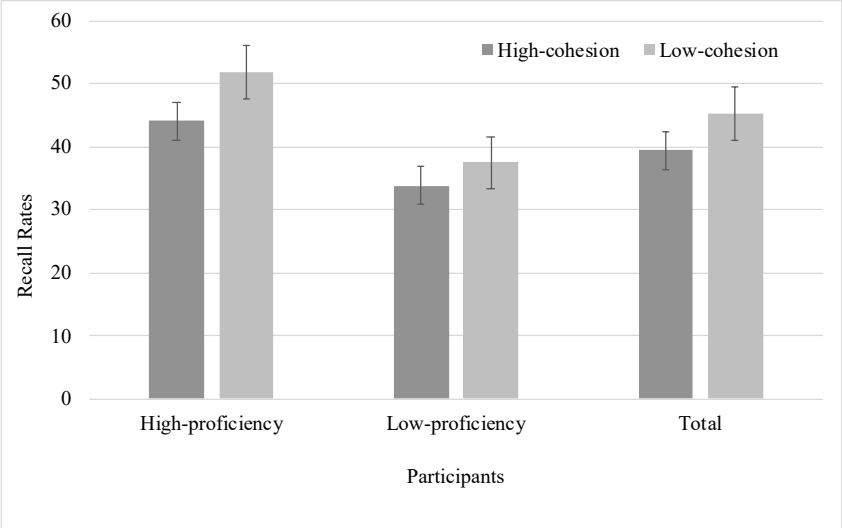
Descriptive Statistics for the Percentage of Recall Production With Arcsine Transformation in Experiment 2-A

		Low-cohesion			High-Cohesion		
Proficiency	*n*	*M*	95%CI	*SD*	*M*	95%CI	*SD*
High	13	51.82	[46.15, 57.48]	9.37	44.13	[39.76, 48.49]	7.23
Low	11	37.51	[31.80, 42.23]	8.51	33.77	[28.42, 39.12]	7.96
Total	24	45.26	[40.44, 50.08]	11.41	39.38	[35.54, 43.22]	9.09

Note. CI = confidence interval; SD = standard deviation.

Figure 4.4

Means of Recall Rates of High- and Low-Cohesion Texts by Readers' Proficiency in Experiment 2-A



Note. Error bars represent the standard errors of the means.

As table 4.10 shows, high-proficiency readers (n = 13) recalled more information (M = 51.82, SD = 9.37) when reading the low-cohesion text, whereas low-proficiency

readers (n = 11) recalled less information about the low-cohesion text compared to the high-proficiency readers (M = 37.51, SD = 8.51). Similarly, when reading the high-cohesion texts, skilled readers recalled more information (M = 44.13, SD = 7.23) than poor readers (M = 33.77, SD = 7.96). Regardless of readers' proficiency, more information was recalled from the low-cohesion texts (M = 45.26, SD = 11.41), whereas less information was reproduced on average from the high-cohesion texts (M = 39.38, SD = 9.09).

A 2 (proficiency: high-proficiency, low-proficiency) × 2 (text: high-cohesion, low-cohesion) two-way mixed ANOVA was conducted on the mean recall rates, with Proficiency as a between-participant variable and Text as a within-participant variable. The results showed a significant main effect of text cohesion, $F(1, 22) = 9.40$, $p = .006$, $\eta^2 = .07$, and the effect size was medium to large. The results also showed a significant main effect for Proficiency, $F(1, 22) = 18.74$, $p < .001$, $\eta^2 = .34$, and the effect size was large. The Proficiency × Text interaction effect was not significant, $F(1, 22) = 1.12$, $p = .301$, $\eta^2 = .01$, and the effect size was small.

These results indicate that high-proficiency readers statistically significantly recalled more information than low-proficiency readers in both reading conditions, and, regardless of readers' proficiency, readers had statistically significantly better recall of the low-cohesion text rather than the high-cohesion text. That is, readers' proficiency and text cohesion were the key to constructing textbase-level mental representation, but these factors did not contribute as an interaction factor.

4.2.1.3.3 Why-Question Task

Table 4.11 shows the number of participants and percentages of the three answer groups, reflecting participants' answer coherence for the why-question task. In the low-cohesion text, 12 participants answered why-questions with low coherence (they got zero or one point for the task), five participants with medium coherence (they got two or three points for the task), and seven participants with high coherence (they got four or five points for the task). In the high-cohesion text, 13 participants answered with less coherence, ten participants answered with middle coherence, and one participant answered with high coherence.

The result of a Fisher's exact test showed that the participants' understanding of the text was significantly related to the text condition (p = .048). This result indicates that readers could construct a situation model of the text when reading the low-cohesion text rather than the high-cohesion text. This suggests that the low-cohesion text facilitated readers' construction of coherent representation compared to the high-cohesion text.

Table 4.11

Understanding High- and Low-Cohesion Texts' in Three Different Coherence Levels of Answers to the Why-Question in Experiment 2-A

	Low		Medium		High	
Text	*n*	%	*n*	%	*n*	%
Low-cohesion	12	50	5	21	7	29
High-cohesion	13	54	10	42	1	4

Note. Each category (low, medium, high) shows the level of understanding of the participants. Low = low coherent answer; Medium = medium coherent answer; High = high coherent answer.

4.2.1.3.4 Questionnaire

Table 4.12 shows the proportion of readers' answers in the questionnaire for each text. For Q1: "Were you able to read the text precisely?", 87.50% of the participants (n = 21/24) felt they could read the text sufficiently well when reading the low-cohesion text, whereas 70.83% (n = 17/24) when reading the high-cohesion text. Regarding Q5: "Did you have problems in comprehending the text?", 16.67% of participants (n = 4/24) felt they had difficulty reading the low-cohesion text, and 29.17% felt the same in reading the high-cohesion text (n = 7/24). For Q3: "Were you able to visualize events and scenes written in the text while you were reading?", 91.67% of the participants (n = 22/24) felt they could visualize the situation in their mind when reading the low-cohesion text, whereas 83.33% of the participants (n = 20/24) thought the same when reading the high-cohesion text.

These results indicate that the participants felt that the low-cohesion text was easier to comprehend than the high-cohesion text. Moreover, the participants felt able to visualize the image of the low-cohesion text compared to the high-cohesion text. These tendencies suggest that low-cohesion texts are beneficial for readers.

Table 4.12

Number and Percentage (%) of Participants' Answered Scales in Questionnaire for Each Text in Experiment 2-A

	Scale 1				Scale 2				Scale 3				Scale 4				Scale 5			
	LC		HC		LC		HC		LC		HC		LC		HC		LC		HC	
Question	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Q1	0	0.00	1	4.17	2	8.33	1	4.17	1	4.17	5	20.83	14	58.33	14	58.33	7	29.17	3	12.50
Q2	0	0.00	0	0.00	0	0.00	1	4.17	3	12.50	3	12.50	13	54.17	9	37.50	8	33.33	11	45.83
Q3	0	0.00	0	0.00	0	0.00	2	8.33	2	8.33	2	8.33	11	45.83	13	54.17	11	45.83	7	29.17
Q4	0	0.00	0	0.00	0	0.00	0	0.00	5	20.83	5	20.83	15	62.50	14	58.33	4	16.67	5	20.83
Q5	1	4.17	3	12.50	16	66.67	11	45.83	3	12.50	3	12.50	4	16.67	6	25.00	0	0.00	1	4.17
Q6	1	4.17	0	0.00	5	20.83	1	4.17	4	16.67	6	25.00	7	29.17	15	62.50	7	29.17	2	8.33
Q7	1	4.17	1	4.17	4	16.67	4	16.67	4	16.67	3	12.50	13	54.17	12	50.00	2	8.33	4	16.67
Q8	0	0.00	0	0.00	1	4.17	1	4.17	1	4.17	3	12.50	12	50.00	11	45.83	10	41.67	9	37.50

Note. LC = low-cohesion; HC = high-cohesion. Each item of questionnaire are showed below: Q1 = "Were you able to read the text precisely?," Q2 = "Was the text easy to read?," Q3 = "Were you able to visualize events and scenes written in the text while you were reading?," Q4 = "Was the text clear and obvious?," Q5 = "Did you have problems comprehending the text?," Q6 = "Did you especially pay attention to detailed information?," Q7 = "Were you able to recall the text content smoothly and answer the questions?," Q8 = "Was the information explained in the text coherent and easy-to-grasp?."

Table 4.13 shows the mean number of participants who answered each Likert scale of the questionnaire for each text. As shown in Table 4.13, high-proficiency readers (n = 13) felt they could understand texts well when reading both types of texts more than low-proficiency readers (n =11) in Question 1 (Q1). According to the results of Q5, low-proficiency readers felt it was more difficult to read the text than high-proficiency readers. In addition, the results of Q3 and Q6 showed that high-proficiency readers thought they used strategies (e.g., visualizing the situation) more than low-proficiency readers did.

Table 4.13

Means With SDs of the Five-Point Scale Scores for Questionnaires by Proficiency and Text Type in Experiment 2-A

	High-Proficiency (n = 13)		Low-Proficiency (n =11)	
Item	Low-Cohesion	High-Cohesion	Low-Cohesion	High-Cohesion
Q1	4.31 (0.48)	3.85 (0.80)	3.82 (1.08)	3.55 (1.04)
Q2	4.23 (0.60)	4.46 (0.88)	4.18 (0.75)	4.00 (0.78)
Q3	4.46 (0.66)	4.31 (0.63)	4.27 (0.65)	3.73 (1.01)
Q4	4.08 (0.64)	4.15 (0.56)	3.82 (0.60)	3.82 (0.75)
Q5	2.15 (0.69)	2.38 (1.12)	2.73 (0.91)	2.91 (1.14)
Q6	3.92 (1.04)	4.08 (0.49)	3.18 (1.40)	3.36 (0.67)
Q7	3.69 (0.95)	3.85 (0.80)	3.18 (1.08)	3.27 (1.35)
Q8	4.46 (0.52)	4.23 (0.93)	4.09 (0.94)	4.09 (0.70)

Note. The numbers in brackets are standard deviation.

A 2 (text: high-cohesion, low-cohesion) × 2 (proficiency: high-proficiency, low-proficiency) × 8 (question: Q1 to Q8), three-way mixed ANOVA was conducted on the

mean questionnaire rates, with Text and Question as within-participant variables, and Proficiency as a between-participant variable. The results showed a significant main effect for Question, $F(7, 154) = 21.32$, $p < .001$, $\eta^2 = .27$, with a large effect size. The results also showed a significant main effect for Proficiency, $F(1, 22) = 4.50$, $p = .045$, $\eta^2 = .02$, with a small effect size. However, the main effect of Text was not significant, $F(1, 22) = 0.09$, $p = .770$, $\eta^2 < .01$. Other interaction effects were also insignificant: Text × Proficiency, $F(1, 22) = 0.09$, $p = .770$, $\eta^2 < .01$, Question × Proficiency, $F(7, 154) = 2.35$, $p = .064$, $\eta^2 = .03$, Text × Question, $F(7, 154) = 1.12$, $p = .349$, $\eta^2 = .01$, Text × Question × Proficiency, $F(7, 154) = 0.31$, $p = .840$, $\eta^2 < .01$.

These results validate the idea that each question assessed a significantly different aspect of reading. High-proficiency readers significantly understood the text better than low-proficiency readers, and low-proficiency readers had more difficulty understanding the text compared with high-proficiency readers. The results also show the difference in reading strategies between high- and low-proficiency readers.

4.2.1.4 Discussion

Response to RQ3-1: ***Does high or low cohesion affect Japanese EFL readers' surface-level comprehension of social expository texts?***

The results of the two-way ANOVA of the written recall task showed that Japanese EFL college students recalled significantly more information when reading the low-cohesion text than when reading the revised elaborative text. This result suggests that text cohesion affects learners' textbase-level comprehension. This is contrary to the intuitive expectation from previous L2 studies suggesting that high-cohesive texts are beneficial (Hosoda, 2016, Yano et al., 1994). Rather, our results show that non-revised low-cohesion text is of greater benefit for L2 learners.

Despite the disagreement with some studies, our findings support the reverse cohesion effect shown in several other studies (McNamara, 2001; McNamara et al., 1996; McNamara & Kintsch, 1996; O'Reilly & McNamara, 2007), that is, low-cohesion texts are good for readers (e.g., McNamara, 2001; McNamara et al., 1996; McNamara et al., 2011). Possible reasons for the results are as follows:

It is possible that participants in this study had higher proficiency than expected, thereby causing a reverse cohesion effect (e.g., O'Reilly & McNamra, 2007). Previous studies (e.g., McNamara, 2001; McNamara et al., 1996; McNamara & Kintsch, 1996) suggested that if readers have sufficient knowledge, then they recall more information from low-cohesion text than high-cohesion text. This suggestion is plausible, given the results of the questionnaire. Participants in this experiment felt that they could understand the low-cohesion text better than the high-cohesion text. In addition, they felt that the high-cohesion text was unclear to understand based on the results of the questionnaire. Given these facts, for undergraduate EFL learners who have a certain level of English proficiency, elaborative text revision would have low potential to facilitate readers'

comprehension.

Moreover, the experimental texts were controlled in terms of the vocabulary levels. Thus, even the low-cohesion text where overlapped expressions and connectives were not used frequently, were deemed easy to read for the college-level participants. Given that previous studies assessing cohesion manipulations on reading focused on difficult science or long historical texts, these experimental texts controlled in terms of readability could be too easy for the participants (e.g., Ozuru et al., 2009; Vidal-Abarca, 2000). Therefore, text revision to increase cohesion did not play a good role in facilitating readers' understanding.

Response to RQ3-2: ***Does high or low cohesion affect Japanese EFL readers' deep-level comprehension of social expository texts?***

The result of the Fisher's exact test for the why-question answers showed that the number of participants who comprehended texts at different levels (i.e., coherent understanding or not) were significantly related to text cohesion (i.e., high-cohesion/low-cohesion). This result indicates that Japanese EFL readers could accomplish a deeper understanding when reading low-cohesion texts compared to high-cohesion texts. Thus, the answer to RQ3-2 is that while text revision in terms of cohesion affects readers' construction of the text representation, the revised high-cohesion text did not show a benefit for readers in this case. This result supports the findings of previous studies, that is, the reverse cohesion effect.

According to past studies, readers who can infer the information to fill in cohesion gaps by their own skills might fall into shallow processing (e.g., word analysis) when they read high-cohesion text (Hosoda, 2016; O'Reilly & McNamara, 2007). This is reasonable because if the high-cohesion text provides enough information for

comprehension, readers do not have to infer the relationships stated implicitly (Ozuru et al., 2009). Thus, a possible explanation for this result is readers' successful inference processing when reading the low-cohesion control text that calls for readers' processing to connect gaps within the text (Radvansky et al., 2014).

Another possible reason for the results is the nature of the why-question task. In this experiment, participants read experimental texts for understanding and completing the written recall task first. Then, they worked on the why-question task. Due to completing this task after the reading exercises, participants' memory and background knowledge could have been used for answering the why-question task. Thus, even in the low-cohesion text, they might be able to answer the task coherently.

However, in this study, there was no on-line method used to assess readers' inference generation. By using the why-question task, readers' situation-model-level understanding was measured in terms of answering causal questions. However, this activity was performed after reading; thus, the answers to the why-question tasks may not entirely convey the readers' situation-model-level comprehension. Hence, although there was an effect of text condition on readers' situation-model-level comprehension, a more detailed method to describe readers' on-line processes (i.e., inference generation) is needed to support this result.

Response to RQ3-3: ***Does the cohesion effect on the comprehension of social expository texts vary as per the readers' proficiency level?***

The results of the written recall task and the why-question task did not show any interaction of Proficiency × Text. Consequently, it was demonstrated that text cohesion did not have a different effect depending on learners' proficiency. This result matches previous studies conducted on Japanese EFL learners (Hosoda, 2016). However, previous

L1 studies (Linderholm et al., 2000; McNamara, 2001) have suggested that the effects of text factors differ depending on learners' proficiency, and knowledge, and text difficulty.

A possible explanation for this result is that the experimental materials were too easy for the participants. Thus, no interaction effects were observed. This result could be possible if the participants were less proficient in reading, and the interaction effects between text cohesion and readers' proficiency on readers' comprehension might be observed. Indeed, previous studies have reported interactional effects (e.g., Linderholm et al., 2000; McNamara, 2001; O'Reilly & McNamara, 2007). Hence, to observe the effect (or possibly an interactional effects) of revising text cohesion on readers, further studies dealing with less proficient Japanese EFL readers are needed.

In summary, the results from Experiment 2-A suggest reverse cohesion effects in expository reading among Japanese EFL readers. However, to test the hypothesis that text cohesion will affect readers' proficiency differently, additional experiments are needed. Therefore, Experiment 2-B was conducted.

4.2.2 Experiment 2-B

4.2.2.1 Purpose and Research Questions

Follow-up experiment 2-B, was conducted to extend Experiment 2-A to less proficient Japanese EFL readers. Again, the goal of Experiment 2 (consisting of Experiments 2-A and 2-B) was to examine the effects of text cohesion on EFL Japanese readers' comprehension of expository texts on social topics. Therefore, almost the same research method and procedure as in Experiment 2-A were used in Experiment 2-B. The only differences from Experiment 2-A were the participants and the comprehension measures to be analyzed. The participants were Japanese university students with a relatively lower level of English proficiency than those in Experiment 2-A. For the comprehension task, only the written recall task was used as a measure of comprehension, not the why-question task, because of the low level of proficiency of the readers and the small amount of data available for analysis. Therefore, the RQs were the same as those in Experiment 2-A.

RQ3-1: Does high or low cohesion affect Japanese EFL readers' surface-level comprehension of social expository texts?

It was predicted that high-cohesion texts would be more effective than low-cohesion texts because the proficiency of the readers was lower than in Experiment 2-A. Alternatively, there might be the possibility that none of the text modifications were effective. This could happen if the high-cohesion text was effective for poor readers and the proficiency of the current experiment was higher than that of the threshold. A follow-up experiment was conducted to investigate this possibility.

4.2.2.2 Method

4.2.2.2.1 Participants

Twelve Japanese EFL undergraduate students participated in Experiment 2-B (12 females; average age = 19.08, range = 19–20). Their native language was Japanese, and they had studied English for at least six years in the Japanese education system. Their majors in university were culture and tourism, English communication, human relations studies, Japanese cultural studies, living space design, and psychology. English proficiency was assumed to be around CEFR A1 based on the self-reported scores of standardized English tests.

4.2.2.2.2 Materials

Experimental texts

The same texts as in Experiment 2-A were used. Text cohesion operations were performed on the two expository texts on social topics, and high-cohesion and low-cohesion versions were provided for each text. See Section 4.2.1.2.2 for further details.

Written recall task

In addition to the experimental texts, the written recall task was the same as in Experiment 2-A. This task was used to measure participants' textbase-level comprehension. This task was conducted in the readers' native language, Japanese. To prevent readers from being unable to write all the information they remembered due to limited time, the task was performed with no time limit.

Why-question task

The same why-question task as in Experiment 2-A was used to measure participants'

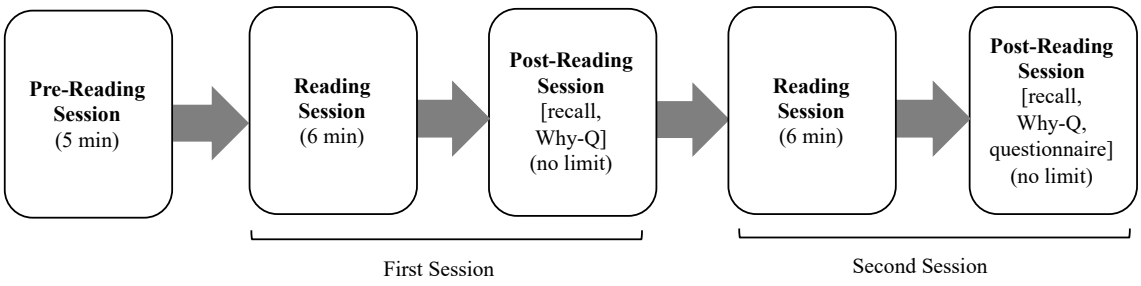
situation-model-level comprehension. However, this task was not analyzed in this experiment because of the low response rate of the participants or the inability of some of the participants to complete the task.

4.2.2.2.3 Procedure

Since Experiment 2-B was a follow-up study of Experiment 2-A, the procedure was based Experiment 2-A (see Figure 4.5). See Section 4.2.1.2.3 for the detailed information. However, this experiment was not conducted individually, but simultaneously for multiple participants. The experiment lasted for approximately 100 minutes. First, participants received an explanation of the study (pre-reading session), then they read two experimental texts in different cohesion versions: one for the low-cohesion condition and one for the high-cohesion (reading session). After reading each text, participants were asked to write down as much text information as possible in Japanese, without referring to the text (post-reading session). In this experiment, the participants also worked on the why-question task as in Experiment 2-A, but because the response rate was low, it was not included in the subsequent analysis.

Figure 4.5

Procedure of Experiment 2-B



Note. recall = written recall task; why-Q = why-question task.

4.2.2.2.4 Scoring and Data Analysis

Written recall task

Written recall was scored using the same procedure as in Experiment 2-A. The written recall rate of the participants was calculated based on the IU segmentation using Ikeno's criteria (Ikeno, 1996) performed in Experiment 2-A. See Section 4.2.1.2.4 for more detailed information. Then, an arcsine transformation was conducted on the total recall production rates because each text had a different number of IUs. The finalized recall production was analyzed using a paired *t*-test to examine the effect of text revisions on textbase-level comprehension.

4.2.2.3 Results

4.2.2.3.1 Written Recall Task

Table 4.14 shows the mean recall rate and standard deviations for each text in the written recall task.

Table 4.14

Descriptive Statistics for the Percentage of Recall Production With Arcsine Transformation in Experiment 2-B

Low-cohesion			High-cohesion		
M	95%CI	*SD*	*M*	95%CI	*SD*
26.36	[22.04, 30.67]	6.79	23.39	[17.89, 28.88]	8.65

Note. CI = confidence interval; SD = standard deviation.

In general, the recall rates of participants in Experiment 2-B were lower than those

in Experiment 2-A (M = 42.32 in Experiment 2-A, M = 23.38 in Experiment 2-B). As Table 4.14 shows, participants in Experiment 2-B recalled more information in the low-cohesion condition (M = 26.36, SD = 6.79) than in the high-cohesion condition (M = 23.39, SD = 8.65).

To check the difference in comprehension between high-cohesion and low-cohesion texts, a paired t-test was conducted on the mean recall rates, with Text as a within-participant variable. The results showed no significant main effect for text cohesion, $t(11) = 1.28$, $p = .227$, $d = .37$, and the effect size was small to medium. This result suggests that relatively low-proficiency readers in Experiment 2-B had the same level of understanding in the high-cohesion and low-cohesion texts.

4.2.2.4 Discussion

Response to RQ3-1: ***Does high or low cohesion affect Japanese EFL readers' surface-level comprehension of social expository texts?***

The results of the t-test of the written recall task showed that relatively low-proficiency Japanese EFL college students recalled the same amount in the high- and low-cohesion texts. No statistically significant difference in the recall amount between high- and low-cohesion texts was observed. This result indicates that text cohesion does not have an impact when learners' proficiency is relatively low. These results are inconsistent with the results of Experiment 2-A. However, considering the difference in the participants' proficiency levels, these inconsistent results make sense.

In Experiment 2-B, relatively low-proficiency university students participated in the experiment. Because they were not skilled readers, there was no reverse cohesion effect (e.g., McNamara, 2001; McNamara et al., 1996; McNamara & Kintsch, 1996; O'Reilly & McNamara, 2007). It is possible that the low-cohesion texts did not promote

comprehension because the proficiency of the readers was not at a sufficient level to allow them to actively generate inferences and achieve deep understanding in the low-cohesion texts.

However, contrary to studies that suggest that high-cohesion texts are effective for less proficient readers (e.g., Bexk et al., 1991; Linderholm et al., 2000; Loxterman et al., 1994), no such effect was found here. This may be explained by the genre of text used in this experiment. In this experiment, expository texts related to social topics were used. Therefore, the texts were more familiar to the readers and relatively easier to read than the scientific texts or technical books used in the previous study (e.g., McNamara et al., 1996), suggesting reverse cohesion effects. Hence, it is possible that some kind of redundant text correction, such as high-cohesion texts, may not be necessary for the comprehension of the texts used in this study. Since the social topic itself was not difficult to read, it is possible that the text modification to make it highly cohesive did not make a significant difference from low-cohesion.

Nevertheless, it is very interesting to note that the effect of cohesion is different for different levels of reader proficiency (see the results of Experiments 2-A, 2-B). Although it was not possible to directly investigate the interaction between cohesion and readers' proficiency on a large scale, this study found that text cohesion in social-topic texts has different effects on comprehension for different levels of proficiency.

4.2.3 Conclusion of Experiment 2 (2-A & 2-B)

Experiment 2 was conducted to explore the effects of text cohesion on EFL Japanese readers' comprehension of expository texts on social topics. Specifically, three RQs were formulated.

RQ3-1: Does high or low cohesion affect Japanese EFL readers' surface-level

comprehension of social expository texts?

RQ3-2: Does high or low cohesion affect Japanese EFL readers' deep-level comprehension of social expository texts?

RQ3-3: Does the cohesion effect on the comprehension of social expository texts vary as per the readers' proficiency level?

To investigate the effects of text revision in terms of cohesion on readers' comprehension of expository texts on social topics, high-cohesion and low-cohesion texts were set, and readers' comprehension was measured by written recall and why-question tasks. Experiment 2 consisted of two experiments targeting Japanese EFL readers of different proficiency levels. In Experiment 2-A, relatively high-proficiency university students participated in the experiment, and relatively low-proficiency university students participated in Experiment 2-B. Based on the results of Experiments 2-A and 2-B, there are some main findings.

First, text cohesion affected readers' literal and deep comprehension (RQ3-1, RQ3-2), In particular, when readers' English proficiency was at a certain level (Experiment 2-A), low-cohesion texts better facilitated readers' comprehension compared to the high-cohesion texts.

Second, based on the results of Experiments 2-A and 2-B, the effects of text revision on text cohesiveness on readers' comprehension differed depending on readers' proficiency levels (RQ3-3). Specifically, when readers had sufficient proficiency, low-cohesion texts improved their comprehension (Experiment 2-A). However, when readers' proficiency was relatively low, text cohesion had no effect on comprehension. Taken together, high-cohesion texts could impede comprehension among EFL readers at a certain proficiency level; low-cohesion texts could facilitate readers' understanding. By

contrast, when readers' proficiency was not very high, cohesion manipulation was not effective; their understanding was not different in the various cohesive texts.

However, the results obtained in this study were limited to reading expository texts on social topics. It is possible that the high-cohesion texts did not promote text comprehension because the content, such as education and the natural environment, was familiar to everyone. Given that previous studies have not sufficiently examined text genre and cohesion, in the following Experiment 3, the author examined whether there was a difference in the comprehension of Japanese EFL learners when reading different cohesive expository texts on science topics.

4.3 Experiment 3: Effects of Text Cohesion on Japanese EFL Readers' Comprehension of Expository Texts on Science Topics

4.3.1 Experiment 3-A

4.3.1.1 Purpose and Research Questions

Experiment 3 aimed to explore the effects of text cohesion on EFL Japanese readers' comprehension of expository texts on science topics. In this experiment, I aimed to verify whether the results obtained in Experiment 2 could be applied to texts of other genres, in particular, science texts, which have been considered difficult to understand in previous studies. The science topics were examined using high-cohesion and low-cohesion texts. Since Experiment 3 focuses on science texts, the cohesion manipulation on text modification for scientific texts was performed following previous studies (McNamara et al., 1996).

Similar to Experiment 2, Experiment 3 consisted of two experiments: Experiment 3-A and Experiment 3-B. Each experiment targeted a different group of university students and Japanese EFL readers with different proficiency levels. The research questions that Experiment 3 attempted to address are as follows:

RQ4-1: Does high or low cohesion affect Japanese EFL readers' surface-level comprehension of scientific expository texts?

RQ4-2: Does high or low cohesion affect Japanese EFL readers' deep-level comprehension of scientific expository texts?

RQ4-3: Does the cohesion effect on the comprehension of scientific expository texts vary as per the readers' proficiency level?

The results indicate that scientific texts are more specialized in content and difficult

to understand than general texts; therefore, high-cohesion texts with heightened relationships and explicit information will help facilitate readers' understanding. However, since Experiment 2 suggested that low-cohesion texts may be effective when the reader's proficiency level is high, such a reverse cohesion effect may be observed here as well. Experiment 3-A targeted students at the same university as in Experiment 2-A. In other words, it has been tested on readers with the same level of proficiency.

It should also be noted the data to be analyzed in Experiment 4 (readers' thinking-aloud protocols) were collected alongside the data of Experiment 3-A. The details of the data were explained in Experiment 4.

4.3.1.2 Method

4.3.1.2.1 Participants

Forty Japanese EFL undergraduate and graduate students participated in this experimental study (25 males and 15 females; average age = 20.90, range = 19–23 years). Their native language was Japanese, and they had studied English for at least six years in the Japanese education system. The participants' majors in universities were varied, including humanities, Japanese language and culture, international studies, social science, education, disability science, biological science, agro-biological resource science, physics, chemistry, policy and planning science, information science, media arts, science and technology, knowledge and library science, health and physical education, medicine, and medical science.

The participants were divided into two reading proficiency groups based on their test scores on an L2 reading proficiency test. The proficiency group was used as a factor in the analysis of the results.

4.3.1.2.2 Materials

Experimental texts

In Experimental Study 3, four expository texts were used (see Appendices A, B, M, and N). The two texts ("Effects of Heat [on Objects, Matter]; EH," "The Needs of Plants; TNP"), which explain scientific ideas, were adopted from McNamara et al. (2011). There were two versions of cohesions for each text (high-cohesion version/low-cohesion version). The two texts were chosen for our study based on three criteria: (a) the text genre was science expository, describing simple scientific phenomena; (b) text manipulation was used in a prior study to increase text cohesion, and (c) the difficulty level and length of the texts were suitable for adult EFL learners.

To further make the texts legible for EFL readers, one paragraph was deleted from each text. Moreover, the difficulty of lexical expressions and low-frequency words of five or higher in the *JACET 8000 List* (JACET Basic Word Revision Committee, 2003) was adjusted by replacing difficult items with simpler words or annotations. A native English speaker then checked the materials for suitability. Table 4.15 presents the primary features of each text.

Table 4.15

Textual Features of Experimental Texts Used in Experiments 3 and 4

	High-cohesion			Low-cohesion		
Text	Words	Sentences	FKGL	Words	Sentences	FKGL
EH	338	27	5.4	264	34	3.2
TNP	359	27	4.0	245	24	3.4

Note. EH = "*Effects of Heat*," TNP = "*The Needs of Plants*," FKGL= Flesch-Kincaid grade level.

Text cohesion manipulation was originally based on the criterion of McNamara et al. (1996). The manipulation aimed to help readers' situation-model construction by increasing the cohesion between concepts and ideas in the text. Thus, the following seven methods were used in the high-cohesion texts: (a) replacing pronouns with noun phrases, (b) adding descriptive elaborations, (c) adding sentence connectives, (d) replacing phrases, (e) adding topic headers, (f) adding theme sentences, and (g) moving or rearranging sentences to increase temporal or referential cohesion (McNamara et al., 2011).

Some examples of McNamara et al. 's (2011) manipulated texts are as follows: in the low-cohesion text, *Plants also need minerals. A mineral is a naturally occurring substance that is neither plant nor animal.*, and in the high-cohesion text *Plants also need minerals. A mineral is not a plant or an animal. Instead, a mineral is a substance in the ground that occurs naturally.* As another example, the low-cohesion text states *Most metals are good conductors. Metal pots are used for cooking. Heat from the stove quickly moves through the metal. The heat warms the food.*, and the high-cohesion text conveys the same content as *Most metals are good conductors. For example, metal pots are used for cooking because heat from the stove quickly moves through the metal pots and the heat from the pot warms the food.*

Table 4.16 lists the characteristics of the texts used in the experiment. These characteristics were calculated using the Coh-Metrix web tool, which provides a diagnosis of text features from various perspectives (e.g., narrativeness, word concreteness, syntax simplicity, and cohesion). Since the Coh-Metrix web tool was used to measure cohesiveness of the texts, the differences between the material versions of the experiments were analyzed with the tool. As mentioned earlier, reference cohesion reflects how texts contain overlapped words and ideas across sentences and the entire text,

while deep cohesion is defined as the degree to which the text contains causal and intentional connectives within the text (McNamara et al., 2014). Based on the text analysis results, the difference between high-cohesion and low-cohesion texts was confirmed in the deep cohesion index.

Table 4.16

Cohesion Indexes of the Experimental Texts Used in Experiments 3 and 4

	High-cohesion		Low-cohesion	
Text	Referential	Deep	Referential	Deep
EH	98%	99%	94%	56%
TNP	99%	99%	99%	89%

Note. EH = "*Effects of Heat*," TNP = "*The Needs of Plants*," Referential = referential cohesion, Deep = deep cohesion.

The text assignments in the experiment were randomly counterbalanced. Participants were provided with one of the following four text assignments: (a) high-cohesion EH text first and low-cohesion TNP text second, (b) low-cohesion EH text first and high-cohesion TNP text second, (c) high-cohesion TNP text first and low-cohesion EH text second, and (d) low-cohesion TNP text first and high-cohesion EH text second. All participants were handed two experimental texts under different text conditions (i.e., high-cohesion/low-cohesion).

In the experiment, they were asked to read and understand the text carefully with no time limit. They were also asked to verbalize whatever thoughts came to mind during reading (think-aloud task). The data of the think-aloud task were recorded using an IC recorder; the readers' think-aloud protocols were analyzed in Experiment 4 as the scope

of this study was readers' comprehension measure through paper and pencil tasks, comparing it with Experiment 3-B.

Written recall task

As with the first experimental study, participants' textbase-level comprehension was measured using a free written recall task. The participants completed the task in Japanese with no time limit.

Why-question task

As in the first experimental study, participants' situation-model-level comprehension was measured through a why-question task. Each text had a single question and readers were asked to explain the causality of the event or phenomenon in the text (e.g., "Why can you warm your cold hands by holding a cup of warm soup?").

Questionnaire

Similar to the second experimental study, questionnaires with a five-point Likert scale (5: "I agree," 4: "I tend to agree," 3: "I am neutral," 2: "I tend to disagree," 1: "I disagree.") were employed to check the participants' impressions of the text and reading strategies that they used during reading. In addition to the question items in Experimental Study 2, a new item (i.e., "Had you already known the idea or mechanism explained in the text before the reading?") was added to the questionnaire to determine the participants' prior knowledge of the text content. The passages used in Experiment 3 were chosen based on the assumption that all participants were familiar with the content.

The instructions for answering the questions were the same as in the first experimental study. Participants were asked to complete the questionnaire by circling

their preferred number (1–5) after reading each experimental text with no time limit.

L2 Reading Proficiency Test

As in the first experimental study, to examine participants' L2 reading proficiency, a 24-item test in a multiple-choice format was adopted from the reading section of the pre-first and second grades of the EIKEN test (Gakken Education Publishing, 2014; Obunsha, 2014). Six items were extracted from the pre-test, and 18 items were extracted from the second grade of the test.

As with the previous experiments, participants were given the test booklet and instructed to answer the test. Additionally, they answered either *Yes* or *No* to the question, "Have you read this test passage before?" when they were done reading the passage. This was to confirm that they had not read this passage before the experiment.

4.3.1.2.3 Procedure

Experiment 3-A was conducted in a session lasting about 100 minutes, including four phases: pre-study session, reading session (first and second), post-reading session (first and second), and test session (see Figure 4.6). The same procedure used in Experiment 2 was followed with some exceptions, as detailed below.

In the pre-study session, the researcher explained the purpose and methods of the experimental study. Students who agreed to participate in the experiment were asked to fill in a profile sheet for demographic information, as in the first experimental study. The items in the profile sheet were name, age, sex, department and grade of university, qualification of English, and experience of studying abroad.

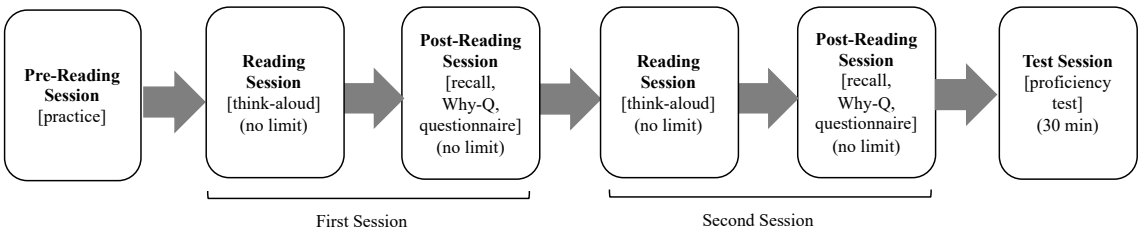
In the reading session, participants read one experimental text within 15.5 minutes and completed the comprehension tasks in the post-reading session. In the latter,

participants answered two comprehension questions (i.e., the written recall task and the why-question task) without referencing the text. Although there was no time limit, they completed the written recall task within 22 minutes, the why-question task within four minutes, and the questionnaire for one experimental text.

The reading and post-reading sessions were repeated in the second reading session using another text with a different revision condition. If a participant had read a high-cohesion text in the first session, they read a low-cohesion text in the second session. Finally, in the test session, the participants took the L2 reading proficiency test, with a time limit of 30 min.

Figure 4.6

Procedure of Experiments 3-A and 4



Note. practice = practice for the think-aloud task; think-aloud = think-aloud task; recall = written recall task; why-Q = why-question task; proficiency test = L2 reading proficiency test.

4.3.1.2.4 Scoring and Data Analysis

Think-aloud protocol

Since Experiment 3-A (and the following Experiment 3-B) focused on effects of text cohesion on reading comprehension, readers' verbal reports collected in the think-aloud method were not scored and analyzed in this experiment.

L2 reading proficiency test

The data were scored by the researcher alone. If a participant's answer was correct, they were given one point for each item. The test consisted of 24 items; thus, the scores ranged from 0 to 24.

Written recall task

The written recall task was conducted in the same way as in Experiment 2. Before scoring, the four experimental texts were divided into IUs (Ikeno, 1996) by two undergraduate students who studied English at university. The inter-rater agreement was 95.61%. Disagreements between the two raters were resolved through discussion. The number of IUs was 54 for EH (low-cohesion) text, 70 for EH (high-cohesion) text, 47 for TNP (low-cohesion) text, and 66 for TNP (high-cohesion) text. Two raters scored 30% of the recall data separately. The inter-rater agreement was 91.44%, and disagreements were resolved through discussion. The researcher scored the remaining data. The recall production rate was calculated after scoring the entire data. Additionally, an arcsine transformation was performed on the total recall production rates because the number of IUs in each text differed.

Why-question task

In Experiment 3-A, why-questions were scored based on four key ideas of the text captured while answering the causal questions (for example, "Why can we warm our hands by holding a cup of warm soup? Explain in Japanese on the basis of the text content that you read before."). Before scoring, two raters decided on three important ideas (e.g., "Heat can move from one object or place to another," "Heat can move from warm objects

to cooler ones," "Heat moves from the soup through the cup to your hands") to answer the causal question and extracted them from the experimental texts. Then, one key idea (e.g., "Heat can move from warm objects to cooler ones.") that had the core information for the answer was chosen from these ideas. If the key idea was expressed in the answer of the participants, one point was given. In addition to the key idea, if participants wrote other ideas, they were given one point for each idea. The maximum possible score was three.

First, the two raters scored 30% of the data separately. The inter-rater agreement was 90.28%, and disagreements were resolved through discussion. The remaining data were scored by the researcher alone. Finally, the sum of the points was calculated. Participants who had the same points were categorized into the same group. Four groups were created, namely, non-coherence group, low-coherence group, medium-coherence group, and high-coherence group. The non-coherence answering group comprised participants who received zero points, the low-coherence group comprised those who got one point, the medium-coherence group comprised those who received two points, and the high-coherence group comprised those who received three points. The number of people in the four groups was compared by text condition (high-cohesion/low-cohesion). To investigate the relationship between readers' comprehension and text cohesion, Fisher's exact test was conducted.

4.3.1.3 Results

4.3.1.3.1 L2 Reading Proficiency Test

The reliability of the L2 reading proficiency test was sufficient (Cronbach's $\alpha = .87$) after excluding one item because of its low discriminability. Before the main analysis, the participants were classified into two reading proficiency groups (see Table 4.17). The

median of the test scores was used to divide participants into a high-proficiency group (n = 17) or a low-proficiency group (n = 23). The mean score of the high-proficiency group (M = 18.53, SD = 2.83) was significantly higher than the low-proficiency group (M = 10.26, SD = 3.14), $t(38) = 8.58$, $p < .001$, $d = 2.75$.

Table 4.17

Means With 95% CIs and SDs of the Proficiency Test Results in Experiments 3-A and 4

Proficiency	*n*	*M*	95%CI	*SD*	*Min*	*Max*
High	17	18.53	[17.07, 19.99]	2.83	15.00	23.00
Low	23	10.26	[8.90, 11.62]	3.14	3.00	14.00
Total	40	13.78	[12.16, 15.40]	5.10	3.00	23.00

Note. CI = confidence interval; SD = standard deviation. Maximum possible score was 24.

4.3.1.3.2 Written Recall Task

Table 4.18 and Figure 4.7 show the mean recall rate and standard deviations for each text in the written recall task. As shown in Table 4.18, high-proficiency readers (n = 17) recalled a larger amount of information when reading both high and low cohesion texts (high cohesion: M = 47.38, SD = 8.63, low cohesion: M = 44.27, SD = 8.34) than low-proficiency readers (n = 23), who recalled lesser amount of information (high cohesion: M = 40.83, SD = 6.63, low cohesion: M = 40.98, SD = 6.15). Regarding the difference in cohesion between texts, high-proficiency readers recalled more information from high-cohesion texts than low-cohesion texts, whereas low-proficiency readers recalled more information for low-cohesion texts than high-cohesion texts. The average readers' recall information amount was larger in high-cohesion texts than in low-cohesion

texts (high cohesion: $M = 44.11$, $SD = 8.14$, low cohesion: $M = 42.63$, $SD = 7.25$).

Table 4.18

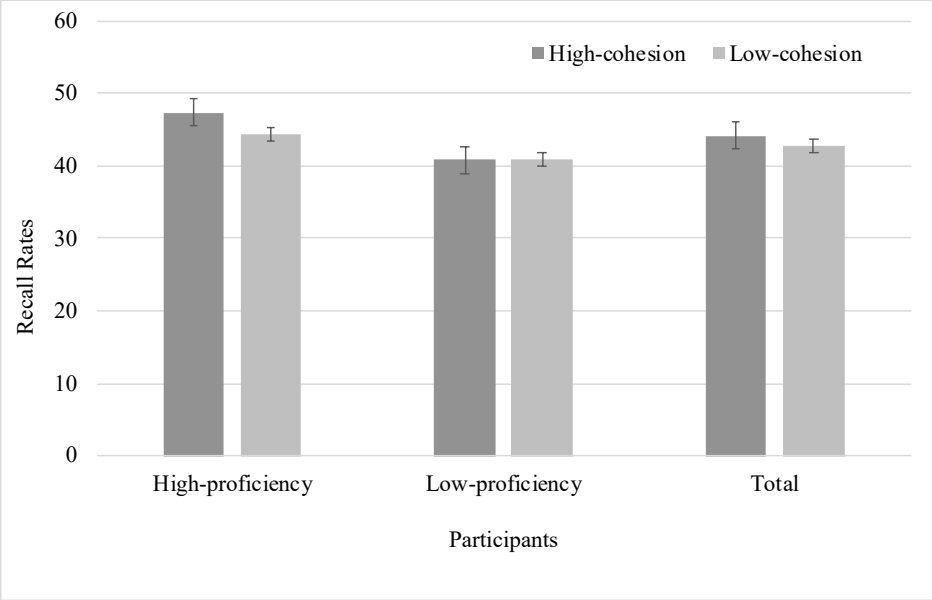
Descriptive Statistics for the Percentage of Recall Production With Arcsine Transformation in Experiment 3-A

		High-cohesion			Low-cohesion		
Prof	*n*	*M*	95%CI	*SD*	*M*	95%CI	*SD*
High	17	47.38	[42.92, 51.84]	8.67	44.27	[39.99, 48.56]	8.34
Low	23	40.83	[37.97, 43.70]	6.63	40.98	[38.32, 43.64]	6.15
Total	40	44.11	[41.10, 46.15]	8.14	42.63	[40.13, 44.63]	7.25

Note. Prof = proficiency; CI = confidence interval.

Figure 4.7

Means of Recall Rates of High- and Low-Cohesion Texts by Readers' Proficiency in Experiment 3-A



Note. Error bars represent the standard errors of the means.

To investigate the interaction between readers' proficiency and text cohesion on comprehension, a 2 (proficiency: high proficiency, low proficiency) × 2 (text: high-cohesion, low-cohesion), two-way mixed ANOVA was conducted on the mean recall rates, with readers' proficiency as a between-participant variable and text as a within-participant variable. The results showed a significant main effect of proficiency, $F(1, 38) = 6.16$, $p = .018$, $\eta^2 = .10$, for which the effect size was medium to large. However, the main effects of Text, $F(1, 38) = 1.36$, $p = .250$, $\eta^2 = .01$, and Proficiency × Text interaction, $F(1, 38) = 1.66$, $p = .206$, $\eta^2 = .01$) were not significant.

These results indicate that high-proficiency readers statistically significantly recalled more information than low-proficiency readers, regardless of text cohesion. The results also indicate that text cohesion does not significantly influence readers' textbase-level understanding of the text.

4.3.1.3.3 Why-Question Task

Table 4.19 shows the number of participants and percentages of the four coherent-comprehension groups (i.e., non, low, middle, and high-coherence comprehension) in the why-question task. The result of Fisher's exact test showed that the participants' understanding at different coherence levels was not significantly related to the text cohesion that participants read ($p = .778$). Thus, text cohesion did not have a significant effect on readers' coherence in the why-question task. The results indicate that participants understood the text and constructed their situation model regardless of the text cohesion. It is, therefore, assumed that readers were able to comprehend the text while reading both high and low-cohesion texts.

Table 4.19

Understanding High- and Low-Cohesion Texts' in Four Different Coherence Levels of Answers to the Why-Question in Experiment 3-A

	Non		Low		Medium		High	
cohesion	*n*	%	*n*	%	*n*	%	*n*	%
High	11	28	2	5	17	43	10	25
Low	8	20	1	3	20	50	11	28

Note. Each category (non, low, medium, or high) shows the level of understanding of the participants. Non = non-coherent answer; Low = low coherent answer; Medium = medium coherent answer; High = high coherent answer.

4.3.1.3.4 Questionnaire

Table 4.20 shows the number of participants and the percentage for each Likert scale answered in the questionnaire for each cohesion text. For Q1: "Were you able to read the text precisely?", 82.50% (n = 33/40) of the participants felt they could understand the text sufficiently when reading a high-cohesion text, and 80% felt the same when reading a low-cohesion text (n = 32/40). For Q5: "Did you have problems in comprehending the text?", 25% (n = 10/40) participants felt they had difficulty reading the high-cohesion text and 35% (n = 14/40) participants had difficulty in reading the low-cohesion text. For Q7: "Were you able to recall the text content smoothly and answer the questions?", 82.50% (n = 33/40) felt they completed the post-reading task smoothly when recalling the high-cohesion text, whereas 60% (n = 24/40) participants were at ease when recalling the low-cohesion text.

These results indicate that the participants felt that the low-cohesion text was easier to comprehend than the revised text. The results also showed that the participants felt they

could complete the written recall and why-question tasks better after they read the low-cohesion text compared to the high-cohesion text. The questionnaire results suggested the benefit of the low-cohesion text for participants' reading and completion of the task.

Table 4.21 shows the means of the five-point Likert scale scores of the questionnaire for each text. Before the analysis, the mean Likert scale for each question was assessed. According to the responses to Q1, regardless of text cohesion, both high-proficiency readers (n = 17) and low-proficiency readers (n = 23) felt they could understand the texts well, but the mean scale was a little lower for low-proficiency readers than for high-proficiency readers. For Q5, low-proficiency readers felt they had difficulty reading low-cohesion text; however, high-proficiency readers experienced difficulty in reading high-cohesion text compared to low-cohesion text. These results indicate a reverse cohesion effect.

Table 4.20

Number and Percentage (%) of Participants' Answered Scales in Questionnaire for Each Text in Experiment 3-A

	Scale 1				Scale 2				Scale 3				Scale 4				Scale 5			
	High		Low		High		Low		High		Low		High		Low		High		Low	
Question	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Q1	0	0.00	1	2.50	5	12.50	2	5.00	2	5.00	5	12.50	19	47.50	22	55.00	14	35.00	10	25.00
Q2	0	0.00	0	0.00	2	5.00	4	10.00	3	7.50	6	15.00	18	45.00	14	35.00	17	42.50	16	40.00
Q3	0	0.00	0	0.00	2	5.00	2	5.00	2	5.00	4	10.00	13	32.50	13	32.50	23	57.50	21	52.50
Q4	0	0.00	0	0.00	2	5.00	4	15.00	7	17.50	3	7.50	20	50.00	22	55.00	11	27.50	11	27.50
Q5	6	15.00	9	22.50	17	42.50	12	30.00	7	17.50	5	12.50	7	17.50	10	25.00	3	7.50	4	10.00
Q6	4	10.00	2	5.00	10	25.00	11	27.50	10	25.00	6	15.00	11	27.50	15	37.50	5	12.50	6	15.00
Q7	2	5.00	1	2.50	3	7.50	9	22.50	2	5.00	6	15.00	24	60.00	14	35.00	9	22.50	10	25.00
Q8	0	0.00	0	0.00	3	7.50	4	10.00	2	5.00	7	17.50	18	45.00	9	22.50	17	42.50	20	50.00
Q9	0	0.00	1	2.50	0	0.00	0	0.00	0	0.00	5	12.50	14	35.00	14	35.00	26	65.00	20	50.00

Note. High = high-cohesion; Low = low-cohesion. Each items' question are as follows: Q1 = "Were you able to read the text precisely?," Q2 = "Was the text easy to read?," Q3 = "Were you able to visualize events and scenes written in the text while you were reading?," Q4 = "Was the text clear and obvious?," Q5 = "Did you have problems in comprehending the text?," Q6 = "Did you especially pay attention to detail information?," Q7 = "Were you able to recall the text content smoothly and answer the questions?," Q8 = "Was the information explained in the text coherent and easy-to-grasp?," Q9 = "Had you already known the idea or mechanism explained in the text before the reading?"

Table 4.21

Five-Point Scale for Questionnaires by Proficiency and Text Cohesion in Experiment Study 3-A

	High-proficiency (*n* = 17)		Low-proficiency (*n* =23)	
Item	High-cohesion	Low-cohesion	High-cohesion	Low-cohesion
Q1	4.12 (1.11)	4.06 (1.14)	4.00 (0.85)	3.87 (0.69)
Q2	4.47 (0.62)	4.35 (0.86)	4.09 (0.90)	3.83 (1.03)
Q3	4.71 (0.59)	4.35 (0.86)	4.22 (0.90)	4.30 (0.87)
Q4	4.29 (0.85)	4.12 (0.78)	3.78 (0.74)	3.91 (0.95)
Q5	2.59 (1.33)	2.41 (1.58)	2.61 (1.08)	2.91 (1.13)
Q6	2.94 (1.44)	3.35 (1.32)	3.17 (1.03)	3.26 (1.10)
Q7	4.18 (0.95)	3.71 (1.21)	3.65 (1.03)	3.48 (1.16)
Q8	4.65 (0.61)	4.35 (1.06)	3.91 (0.90)	3.96 (1.02)
Q9	4.71 (0.47)	4.41 (0.62)	4.61 (0.50)	4.22 (1.04)

Note. Standard deviations are in parentheses. Q1 = “Were you able to read the text precisely?,” Q2 = “Was the text easy to read?,” Q3 = “Were you able to visualize events and scenes written in the text while you were reading?,” Q4 = “Was the text clear and obvious?,” Q5 = “Did you have problems in comprehending the text?,” Q6 = “Did you especially pay attention to detail information?,” Q7 = “Were you able to recall the text content smoothly and answer the questions?,” Q8 = “Was the information explained in the text coherent and easy-to-grasp?,” Q9 = “Had you already known the idea or mechanism explained in the text before the reading?”

Finally, to examine the effect of text cohesion on the readers' impressions of the text and strategy use, a 2 (proficiency: high proficiency, low proficiency) × 2 (text: high cohesion, low cohesion) × 9 (question: Q1 to Q9) three-way mixed ANOVA was conducted on the mean score of the questionnaire, with text and question as within-participant variables, and proficiency as a between-participant variable. The results showed a significant main effect of the questions, $F(8, 304) = 26.76$, $p < .001$, $\eta^2 = .25$, with a large effect size. A small effect size was identified in the main effect of proficiency, $F(1, 38) = 8.61$, $p = .067$, $\eta^2 = .01$. However, the main effect of text was not significant, $F(1, 38) = 1.35$, $p = .252$, $\eta^2 < .01$, and other interactional effects were insignificant too: Text × Proficiency, $F(1, 38) = 0.60$, $p = .442$, $\eta^2 < .01$, Question × Proficiency, $F(8, 304) = 1.24$, $p = .293$, $\eta^2 = .01$, Text × Question, $F(8, 304) = 1.08$, $p = .366$, $\eta^2 = .01$, Text × Question × Proficiency, $F(8, 304) = 0.69$, $p = .584$, $\eta^2 < .01$. The results indicate the validity of the question items in that each question assessed different aspects of reading. The results also suggest differences in reading strategies between high and low-proficiency readers.

Through several analyses of the questionnaire, it was revealed that about 80 % of participants were able to read the text precisely, with more difficulty reading the low-cohesion text. The results also showed that high-proficiency readers experienced more difficulty in reading the high-cohesion text, whereas low-proficiency readers felt it was difficult to read the low-cohesion text. These results suggest the existence of a reverse cohesion effect on Japanese EFL readers.

4.3.1.4 Discussion

Response to RQ4-1: ***Does high or low cohesion affect Japanese EFL readers' surface-level comprehension of scientific expository texts?***

The result of the two-way ANOVA of the written recall task showed that Japanese EFL learners did not recall significantly different amounts of information when reading texts of different cohesion levels. This suggests that differences in text cohesion do not affect learners' textbase-level comprehension. Therefore, the results of Experiment 3-A contradict the results of Experiment 2 in this dissertation and other past studies (Hosoda, 2016; Loxterman et al., 1994; McNamara, 2001; McNamara et al., 1996; McNamara & Kintsch, 1996; O'Reilly & McNamara, 2007; Yano et al., 1994). Two explanations for this result are provided below.

First, there could be a possibility that the experimental texts were not too difficult to read for EFL university students who participated in Experiment 3-A. If a text provides coherent information, no additional revision is needed. This is clear from the fact that prior studies have revised low-cohesion science texts to improve readers' understanding (Hall et al., 2014; McNamara et al., 1996; Ozuru et al., 2010). If this is true, manipulating texts does not necessarily facilitate readers' comprehension. In a study by Linderholm et al. (2000), readers' comprehension is influenced by text cohesion only when reading difficult texts. The experimental texts used in this study were not very difficult, and consequently, the revised high-cohesion text may not have worked effectively. However, in Experiment 3-A, no effects of low-cohesion text were confirmed (i.e., the reverse cohesion effect).

In Experiment 2-B, neither high-cohesion nor low-cohesion texts affected comprehension, possibly due to a threshold of proficiency that does not require either high-cohesion or low-cohesion. As such, the level of proficiency of the learners in

Experiment 3-A matched that of the experimental texts because the difficulty level was controlled, and therefore, none of the text cohesion revisions had any effect.

Second, it could be that the text content was sufficiently well-known for the participants in Experiment 3-A, and they used their own knowledge to understand the texts without using the textual signals, which were increased by cohesion manipulation, to integrate the ideas in the text. A previous L1 study (McNamara et al., 2011) was used to understand the text cohesion effect in comprehending science texts. However, the L1 study was conducted on nine-year-old children, with the same content as science textbooks for elementary or middle school in Japan. The questionnaire results indicated that almost all participants knew the ideas and phenomena explained in the texts. Thus, it is likely that readers relied on their pre-existing knowledge to comprehend the text. Therefore, there were no cohesion effects on the participants' literal understanding of the texts in this experiment.

Response to RQ4-2: ***Does high or low cohesion affect Japanese EFL readers' deep-level comprehension of scientific expository texts?***

The result of Fisher's exact test for the why-question task showed that Japanese EFL readers can accomplish situation-model-level understanding, regardless of text cohesion. Similar to RQ 4-1, this result is inconsistent with prior studies (Hosoda, 2016; O'Reilly & McNamara, 2007) and the results of Experiment 2, in which expository texts on social topics were presented. The possibilities mentioned below may explain this result.

First, it may be that experimental texts were not too difficult for the participants in this experiment to understand, and supplemental aid (high-cohesion revision and/or low-cohesion revision) was not necessary. If this is the case, no effects of text cohesion are visible.

Second, the result could be explained if readers' knowledge strongly influenced their reading, and textual differences in terms of cohesion did not affect their situational comprehension. If readers have sufficient knowledge, they can infer the gaps between ideas and gaps in the text by taking advantage of their own knowledge (O'Reilly & McNamara, 2007). Indeed, some answers written by participants in the why-question task were not stated in the experimental texts. This means that some participants relied on information outside the experimental text or their own knowledge to answer the why-question.

It should be noted that issues regarding text readability and readers' background knowledge imperiled the validity of the why-question task. Therefore, it is possible that the effect of text cohesion was not measured properly since the task was accomplished with the help of the readers' previous knowledge, rather than a coherent understanding of the information stated in the text, which is influenced by the cohesion of the text. Thus, it is difficult to draw conclusions on the readers' consistent understanding of this task alone. The lack of a cohesion effect, similar to the results for the written recall task (RQ4-1), may be an observed consequence of the interactive effects of the participants (relatively high-proficiency readers) and text topics (science texts) in Experiment 3-A.

Response to RQ4-3: ***Does the cohesion effect on the comprehension of scientific expository texts vary as per the readers' proficiency level?***

The results of the text comprehension task based on the reader's proficiency did not confirm an interaction effect between the reader's proficiency and text cohesion. In Experiment 3-A, no statistically significant effect of text cohesion modification was found and no interaction between cohesion and reader proficiency was recorded.

This result is consistent with the results of Experiment 2, which was conducted on

a social topic. In Experiment 2-B, there was no effect of text cohesion, but in the Experiment 2-A conducted with high-proficiency readers, the text cohesion effect was confirmed, suggesting text cohesion can affect differently by readers' proficiency level. Therefore, it is possible that in the present study, due to the lack of variation in proficiency, the interaction effect between the reader's proficiency and text cohesion was not observed. To verify this, we conducted a follow-up experiment to Experiment 3-A (Experiment 3-B) in the same way as Experiment 2-A and 2-B.

4.3.2 Experiment 3-B

4.3.2.1 Purpose and Research Questions

The purpose of Experiment 3-B as a follow-up experiment was to extend the results of Experiment 3-A to low-proficiency Japanese EFL readers. The goal of Experiment 3 (Experiments 3-A and 3-B) was to examine the effects of text cohesion on EFL Japanese readers' comprehension of expository texts on science topics. Almost the same research method and procedure used in Experiment 3-A were employed in Experiment 3-B. The only differences were the participants and comprehension measures to be analyzed. Experiment 3-B targeted Japanese university students with a relatively lower level of English proficiency than Experiment 3-A. For the comprehension task, only the written recall task was used as a measure of comprehension, and not the why-question task, because of the low level of proficiency of the readers and limited data available for analysis.

Thus, the research questions of Experiment 3-B were the same as Experiment 3-A.

RQ4-1: Does high or low cohesion affect Japanese EFL readers' surface-level comprehension of scientific expository texts?

In Experiment 3-B, it was predicted that cohesion manipulation could affect comprehension when readers' proficiency was lower than that of Experiment 3-A. Thus, high-cohesion texts could facilitate low-proficiency readers' understanding, and low-cohesion texts could facilitate high-proficiency readers' comprehension. This could be because the participants in Experiment 3-B have low proficiency and hence, poor understanding, and only high-cohesion texts lead to a good comprehension of readers. A follow-up experiment was conducted to investigate the above possibilities.

4.3.2.2 Method

4.3.2.2.1 Participants

Experiment 3-B was conducted on 20 Japanese undergraduates (aged 19–20 years, M = 19.11, SD = 0.32), majoring in psychology, humanities, culture, and design. All participants were women and native Japanese speakers who had learned English as a foreign language for over six years. Because of incomplete responses (one text out of two for each), data of two participants were excluded. Participants' general English proficiency was estimated to be at the basic to independent level (A1 to A2), per the Common European Framework of Reference for Language based on their self-reported standardized test scores.

4.3.2.2.2 Materials

Experimental texts

Since Experiment 3-B was conducted to collect follow-up data, the same materials used in Experiment 3-A were used. Two versions of scientific expository texts were created for high cohesion and low cohesion. Refer to section 4.3.1.2.2 and Appendices A, B, M, and N for detailed information on the same.

Written recall task

Similar to the experimental text, the comprehension measurement task was adapted from Experiment 3-A. The written recall task, which asks participants to recall and write as much text content as possible after reading, without referring to the text, was used to measure participants' text comprehension. Participants were notified of this recall task before reading the experimental texts. The task was conducted in Japanese with no time limit.

L2 reading proficiency test

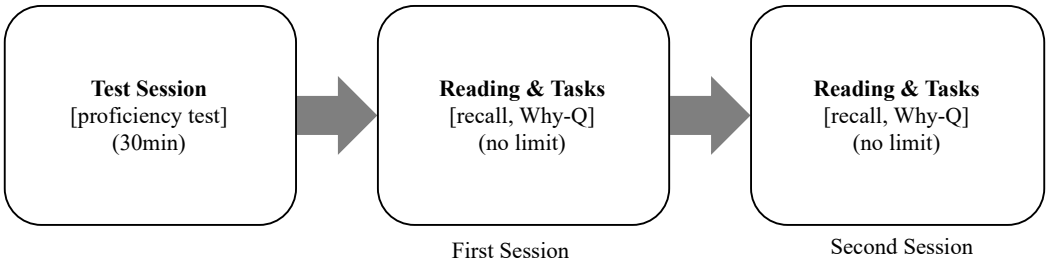
To calculate participants' reading proficiency in English, an L2 reading proficiency test was conducted. A 26-item reading test was adopted from the standardized English proficiency test of Japan (EIKEN test). The test consisted of six passages, two each for pre-second, second, and pre-first grade. The test had a 30-min time limit.

4.3.2.2.3 Procedure

Before the experiment, participants received an explanation of the study procedure. The experiment consisted of two sessions. In the first session, the participants took the L2 reading proficiency test within the 30-min time limit. In the second session, they completed the reading tasks in approximately 60 min. They were then asked to read two texts carefully to complete the post-reading tasks. First, they read one of the experimental texts and completed the written recall task. The same procedure was repeated for the other texts. The order of text assignments was randomly counterbalanced. Specifically, participants either read a high or low-cohesion text in the first assignment and the other in the second (see Figure 4.8).

Figure 4.8

Procedure of Experiment 3-B



Note. proficiency test = L2 reading proficiency test; recall = written recall task; why-Q = why-question task.

4.3.2.2.4 Scoring and Data Analysis

Written Recall Task

The present study replicated the same data analysis procedure as Experiments 2 and 3-A. Before scoring, two raters divided the four experimental texts into idea units (IUs) based on the criteria of Ikeno (1996). The inter-rater agreement for the IU division (95.61%, Cohen's $\kappa = .87$) was confirmed in Experiment 3-A. The experimental texts had the following IUs: 54 for the low-cohesion EH text, 70 for the high-cohesion EH text, 47 for the low-cohesion TNP text, and 66 for the high-cohesion TNP text. If at least two-thirds of the content in an IU was reproduced in the participants' answers, they received one point for that IU. Finally, the recall production rate is calculated.

The R (version 3.6.2, R Core Team, 2018) was used for statistical analyses. To statistically predict the effects of text cohesion on recall rates, a generalized linear mixed model (GLM) using the "lme4" package (Bates et al., 2019) was used because of the small sample size and normality. Effect coding (-0.5 = low-cohesion, 0.5 = high cohesion) was applied to the models as a fixed effect: cohesion. The models also included random intercepts of the subjects and items (see Appendix O for the R scripts).

L2 Reading Proficiency Test

If the participants answered an item correctly, they received one point. The maximum possible score for the test was 26.

4.3.2.3 Results

4.3.2.3.1 L2 Reading Proficiency Test

The descriptive statistics of the L2 reading proficiency test were as follows: M = 6.94, SD = 2.58, Min = 3, Max = 11. The participants were relatively low-proficiency EFL readers compared to Experiment 3-A in both the reading proficiency test and their general English proficiency level.

4.3.2.3.2 Written Recall Task

Table 4.22 shows the descriptive statistics of the mean recall rate for each text. The average recall rate for the high-cohesion texts was greater than that for the low-cohesion texts (high-cohesion text: M = 17.83, SD = 7.67; low-cohesion text: M = 14.83, SD = 10.36). The results of the GLM showed a significant effect of cohesion (estimate = 0.36, SE = 0.14, z = 2.63, p = .008). Figure 4.9 shows the production rate of the written recall task. The shaded area in the figure shows the possible range of the readers' recall rates predicted from the data, and the solid line represents the regression line.

Table 4.22

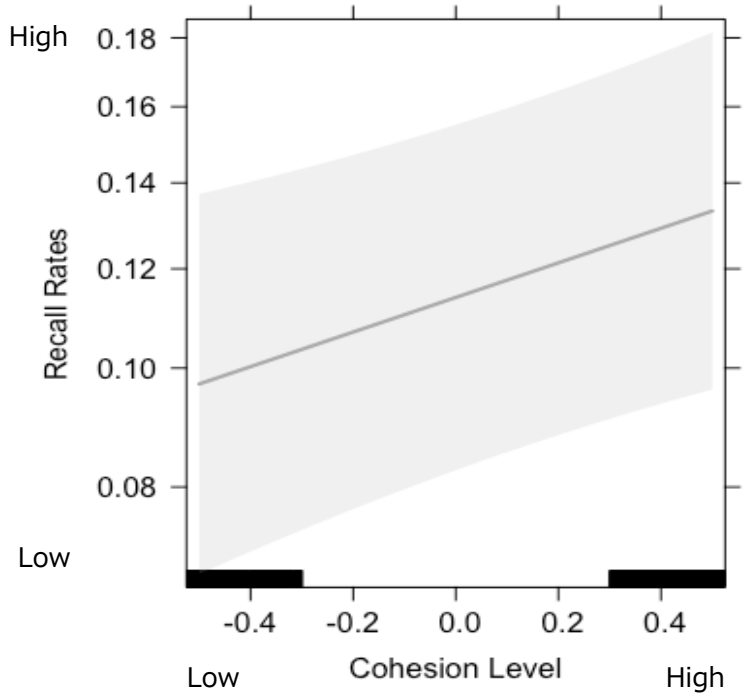
Descriptive Statistics of Participants' Recall Rates for Each Cohesion Text in Experiment 3-B

Text	*n*	*M*	95%CI	*SD*
High-Cohesion	16	17.83	[13.74, 21.92]	7.64
Low-Cohesion	18	14.87	[9.71, 20.02]	10.36
Total	34	16.26	[13.05, 19.46]	9.18

Note. Because of the deviated responses from the task instruction, two participants' data of high-cohesion text were excluded.

Figure 4.9

Participants' Recall Rates Related to Text Cohesion in Experiment 3-B



4.3.2.4 Discussion

RQ4-1: ***Does high or low cohesion affect Japanese EFL readers' surface-level comprehension of scientific expository texts?***

The results of the written recall test suggested that high-cohesion texts facilitated the understanding of relatively poor readers. This finding is reasonable given that informative high-cohesion texts, which eliminate readers' need to guess the missing information owing to gaps in the text, have positive effects on low-proficiency readers' comprehension (Gilabert et al., 2005). A similar cohesion effect has been observed in previous L2 research, such as Hosoda's (2016) results of the free-written recall task and other text revision studies (e.g., Yano et al., 1994). Hosoda discussed in his paper that high-cohesion texts increase opportunities for readers to comprehend the required

information and help them retain and encode text information.

However, the finding that high-cohesion texts facilitate reader comprehension more than low-cohesion texts is not consistent with Experiment 3-A, which was conducted with relatively proficient readers in a similar design. Nonetheless, given that the experimental texts in Experiment 3 were science texts, which can be difficult for readers with limited knowledge and ability, the differences between the experimental results can be reasonably explained. In summary, based on the experimental results so far, high-cohesion texts are effective when the reader's proficiency level is low, and low-cohesion texts are more beneficial when the reader's proficiency level is high. These findings support the existence of reverse cohesion effects in L2 reading similar to L1 studies (e.g., O'Reilly & McNamara, 2007).

In the L1 study, the reverse cohesion effect was found in the differences in background knowledge of the readers; however, in the L2 study, a similar effect may be found in the differences in reading comprehension skills considering L2 readers' limited cognitive resources to comprehend low-cohesion texts (Horiba, 1996). Comparatively, L1 readers do not face as many difficulties as L2 readers, due to the former's background knowledge and ability to make inferences. Therefore, reverse cohesion effects can be applied to L2 reading.

Moreover, the experimental studies conducted in this dissertation suggest that cohesion effects are not applicable in all cases, and there may be a threshold of proficiency at which none of the text cohesion modifications are effective. It also seems that the effectiveness of cohesion revision depends on whether the reader faces difficulty in reading or not, whether the topic is unfamiliar, and whether the reader has prior knowledge of the topic. Thus, the answer to the RQ, *Does high or low cohesion affect Japanese EFL readers' surface-level comprehension of scientific expository texts?* is

affirmative. For low-proficiency readers, high-cohesion texts are a beneficial option when reading unfamiliar scientific text content.

However, it must be acknowledged that this study is limited by the fact that it was a follow-up experiment with limited data and that the only measure of comprehension was the written recall test data. Thus, further studies are needed to investigate text cohesion effects at different levels of readers' comprehension.

4.3.3 Conclusion of Experiment 3 (3-A & 3-B)

Experiment 3 was conducted to explore the effects of text cohesion on EFL Japanese readers' comprehension of expository texts on science topics. Specifically, three research questions emerged as follows:

RQ4-1: Does high or low cohesion affect Japanese EFL readers' surface-level comprehension of scientific expository texts?

RQ4-2: Does high or low cohesion affect Japanese EFL readers' deep-level comprehension of scientific expository texts?

RQ4-3: Does the cohesion effect on the comprehension of scientific expository texts vary as per the readers' proficiency level?

To address the above questions, high-cohesion and low-cohesion texts on scientific phenomena were used, and readers' comprehension was measured by the written recall and why-question tasks. Experiment 3 consisted of two experimental studies, Experiment 3-A and 3-B, similar to Experiment 2. Experiment 3-A targeted relatively high-proficiency university students, and Experiment 3-B, a follow-up study to Experiment 3-A, focused on relatively low-proficiency university students.

Based on the results of Experiments 3-A and 3-B, here are some pertinent findings.

First, regarding the effects of text cohesion revision on Japanese EFL readers' literal comprehension (RQ4-1), high-cohesion text was found to facilitate the comprehension of readers with relatively low proficiency as compared to low-cohesion text. Conversely, when the proficiency level of the readers was relatively high, text cohesion modification did not affect comprehension. These results are consistent with the findings of previous studies that agree that high-cohesion texts help readers comprehend the content better. Interestingly, there were also cases where no effects of high or low cohesion were observed, based on the proficiency of the reader.

Second, regarding Japanese EFL readers' deep comprehension (RQ4-2), the findings suggested that there was no effect of cohesion on comprehension at the situation model level for readers with relatively high proficiency. As the readers were highly proficient and may have had previous knowledge of the text topic, it was suggested that cohesion manipulation may not affect readers' deep comprehension. However, since this finding could be verified with Experiment 3-A, further research is required in the future.

Finally, in terms of interaction effects of text cohesion and readers' proficiency in comprehension of cohesive texts (RQ4-3), no such interactional effects were found in each experiment. However, the results of the two experimental studies that targeted Japanese EFL readers at different proficiency levels indicate interactional effects of text cohesion and readers' proficiency on readers' literal comprehension. Specifically, when readers' proficiency is relatively low, high-cohesion texts can improve their understanding of science texts compared to low-cohesion texts. However, when readers' proficiency is sufficient, neither high nor low-cohesion texts show any facilitating effects on readers' comprehension.

While a reverse cohesion effect was observed in Experiment 2, which used social topics familiar to readers, it was interesting to observe that high-cohesion texts promoted

comprehension among less-proficient readers of science texts, which are likely to be less familiar to readers. In summary, the results revealed that in L2 reading comprehension, both the effect of high-cohesion texts on text comprehension and the effect of low-cohesion texts on text comprehension are observed, based on the topic of the text and the proficiency level of the reader. This is an important finding that should be implemented in education for EFL learners, considering variations in readers' proficiency and L2 experience in the EFL context.

The experimental texts used in Study 2 focused on the reader's understanding of the low and high-cohesion texts after reading them. However, the study was unable to examine the cognitive level of the reader while reading texts of varied cohesion. To resolve such issues, the final experimental study, Experiment 4 aimed to examine whether there was a difference in Japanese EFL learners' cognitive processing during reading different cohesive expository texts on science topics.

4.4 Experiment 4: Effects of Text Cohesion on Japanese EFL Readers' Inference Generation

4.4.1. Purpose and Research Questions

Experiment 4 was designed to further investigate the effects of text revision based on the results of Experiments 2 and 3, where readers' comprehension was investigated. Experiment 4 overcame the limitation faced in Experiments 2 and 3, which measured readers' comprehension only on the basis of post-reading tasks such as written recall, why-questions, and questionnaires, without measuring their cognitive processes or strategies during reading. The offline tasks were only able to indicate the readers' memory and recall outcomes. In contrast, an online task could be used to assess readers' cognitive processing during reading. Thus, to further investigate the effect of text revision on readers' cognitive processes and comprehension, an online task (e.g., think-aloud task) is necessary.

To overcome the limitations discussed above, Experiment 4 focused on the effects of text cohesion on Japanese EFL readers' cognitive processing during reading. To observe the cognitive processing, a think-aloud task was adopted, in which the reader was asked to verbalize what they were thinking while reading. As mentioned earlier, Experiment 3-A employed this method. However, due to the scope of Experiments 3-A and 3-B (effects of text cohesion on comprehension outcome), Experiment 3-A did not analyze the protocol data. Thus, this section (Experiment 4) used the think-aloud protocol data collected in Experiment 3-A.

In addition, the researcher attempted to include not only the language proficiency of the readers, but also their cognitive-allocation tendencies during reading (i.e., whether they tend to generate inferences frequently) as analysis factors. Thus, the main data of this experiment were collected in Experiment 3-A, the goal of the study and analytical

perspective being different from Experiment 3-A. The three research questions in Experiment 4 are given below:

RQ5-1: Does cohesion affect inference generation among Japanese EFL readers while reading scientific expository texts?

RQ5-2: Does the cohesion effect on inference generation from scientific expository texts vary as per the readers' proficiency?

RQ5-3: Does the cohesion effect on inference generation from scientific expository texts vary depending on the readers' processing tendency during reading?

Since the cohesion manipulations help the reader understand science texts (positive effects of high-cohesion texts in Experiment 3-B), it is possible to note differences in comprehension and cognitive processes when the cohesion of the texts is different. Moreover, the type of text facilitating readers' inferences would depend on the proficiency of the readers, as seen in previous experiments (e.g., Horiba, 1996). Specifically, when the reader's proficiency level is high, the high-cohesion text may become redundant and prevent the reader from generating inferences. Therefore, it is expected that drawing inferences and higher cognitive processes will be more active in low-cohesion texts. However, if the reader's comprehension skill is poor and he or she is not able to draw inferences due to weak cognitive processes, inferences may be more easily triggered if the relationships are made explicit in high-cohesion texts.

4.4.2 Method

4.4.2.1 Participants

Since data collection was conducted in Experiment 3-A, the participants of Experiment 4 were the same as those of Experiment 3-A (i.e., 40 Japanese EFL undergraduate and graduate students whose native language was Japanese). The participants were divided into two reading proficiency groups to obtain the reading proficiency test scores in the analysis of Experiment 3-A. Participants were further divided into different reading process groups based on the results of a cluster analysis for the think-aloud protocol. Readers' proficiency and cognitive processing patterns were treated as reader factors in the data analysis.

4.4.2.2 Materials

Experimental texts

The experimental texts were science texts used in Experiment 3-A. The two texts ("Effects of Heat [on Objects, Matter]; EH," "The Needs of Plants; TNP"), which explain scientific ideas, were adopted from McNamara et al. (2011) with two versions of cohesion for each text (high-cohesion/low-cohesion). See Section 4.3.1.2.2 and Appendices A, B, M, and N for further details.

Think-aloud task

Readers' cognitive processing and strategies of reading were assessed through an online think-aloud task. Participants were asked to verbalize their thoughts while reading. They practiced the think-aloud task before the experiment. A model passage extracted from McNamara et al. (1996) was used for practice. (See Appendix P).

In the experiment, participants received a brief instruction of the task, being asked

to read the model passage while listening to a model CD of a think-aloud protocol. After listening to the model CD, they read out short model passages to practice reading while expressing what they were thinking. If they faced further issues during the think-aloud task, the researcher guided them accordingly. Then, they read experimental texts in different cohesion conditions while doing the think-aloud task.

L2 reading proficiency test

As in Experiment 3-A, to examine participants' L2 reading proficiency, a 24-item test was adopted from the reading section of the pre-first and second grades of the EIKEN test (Gakken Education Publishing, 2014; Obunsha, 2014). Six items were extracted from the pre-test, and 18 items were extracted from the second grade of the test.

4.4.2.3 Procedure

Since the data analyzed in Experiment 4 were collected in Experiment 3-A, the procedure followed in this study was the same as Experiment 3-A. The study was conducted in a 100-min session, including four phases: pre-study session, reading session (first and second), post-reading session, and test session (see Figure 4.6 in Section 4.3.1.2.3).

4.4.2.4 Scoring and Data Analysis

L2 reading proficiency test

The data were scored by the researcher alone. If a participant's answer was correct, they were given one point for each item. The test consisted of 24 items; thus, the possible score ranged from 0 to 24.

Think-aloud protocol

To score the participants' think-aloud protocol, this study applied the coding system from frameworks of previous studies (Horiba, 2013; Kimura, 2015; Shimizu, 2015). The frameworks were chosen because the studies focused on Japanese EFL learners' reading process and strategy use, similar to the present study.

First, to determine the framework of this study, some categories used in prior studies were combined or deleted. All participants' audio data were transcribed and divided into almost-clauses by the researcher alone. Then, each clause was categorized into one of the following 12 subcategories: (a) *word analysis*, (b) *sentence analysi*s, (c) *paraphrase*, (d) *backward inference*, (e) *predictive inference*, (f) *association*, (g) *evaluation*, (h) *reaction*, (i) *self-monitoring*, (j) *text organization*, (k) *reading aloud*, and (l) *others* (see Table 4.23).

Table 4.23

Categories of Think-Aloud Protocols Used in Experiment 4

PL	Analysis	Para-phrase	Inference	Response	Meta comment
C	(a) word (b) sentence (k) reading aloud	(c) para-phrase	(d) backward (e) predictive	(f) association (g) evaluation (h) reaction	(i) self-monitoring (j) text organization

Note. PL = processing level; C = categories.

Second, these 12 categories were integrated into five process categories: (1) *analysis*, (2) *paraphrase*, (3) *inference*, (4) *response*, and (5) *meta comment*. These categories and examples of the think-aloud protocol are available in Appendix Q.

Word and *sentence analysis* were combined in the *analysis* category, as these categories included participants' analysis of the form and meaning of text surface

information. The *paraphrase* category was the same as the subcategory of *paraphrase*, in which participants attempted to rephrase the expression of the text in their own words. *Backward inference* and *predictive inference* were integrated into the *inference* process, where the participants tried to understand the ideas in the text and integrated the text content with their own knowledge (e.g., background knowledge). The inference activity leads to a deeper understanding and coherent representation of the text. On the other hand, *response,* including the *association*, *evaluation*, and *reaction* subcategories, was categorized as a process in which participants create an active response based on text content or their background knowledge. The reaction process is distinguished from the inference process in that it does not promote a deeper understanding of the text. The reaction process consists of readers' comments that are not directly related to the text content and can be regarded as a strategy to actively understand the writer's message by analyzing the relationship between the writer, the text, and the reader themselves (Kimura, 2015). *The meta-comment* consisted of two subcategories: *self-monitoring* and *text organization*. This is the process by which participants analyze their understanding of the text from a higher viewpoint.

Two undergraduate students majoring in English categorized 30% of participants' protocols into 12 subcategories. The reliability of raters' categorization was 90.60%. Disagreements were resolved through raters' discussions. The remaining 70% of the data were sorted by the researcher alone. After scoring, the frequency of each category was calculated based on the scores. Additionally, as each participant's comments differed, an arcsine transformation was performed on the total frequency of the five processes' comments and compared from the same perspective.

4.4.3 Results

4.4.3.1 L2 Reading Proficiency Test

Again, the data collected in Experiment 3-A was used in this analysis. As mentioned earlier, the reliability of the L2 reading proficiency test was sufficient (Cronbach's α = .87) after excluding one item because of its low discriminability. Before the main analysis, the participants were classified into two reading proficiency groups (see Table 4.17 in Section 4.3.1.3.1). The median of the test scores was used to divide participants into a high-proficiency group (n = 17) and a low-proficiency group (n = 23). The mean score of the high-proficiency group (M = 18.53, SD = 2.83) was significantly higher than the low-proficiency group (M = 10.26, SD = 3.14), $t(38) = 8.58$, $p < .001$, d = 2.75.

4.4.3.2 Think-Aloud Protocol

Table 4.24 shows the production rate of think-aloud comments of participants per category after reading the content of different cohesion texts. The data were a result of the arcsine transformation, which was conducted because the comments differed depending on the categories or text cohesion. As shown in Table 4.24, the participants frequently produced lower-level processing comments (e.g., analysis) and few comments on higher-level processes (e.g., inferences, responses, and meta comments).

Table 4.24

Proportions of Think-Aloud Comments (%) for Process Level by Text Cohesion in Experiment 4

		High-cohesion			Low-cohesion		
Process	Category	*M*	95% CI	*SD*	*M*	95% CI	*SD*
Analysis	Word	35.35	[32.82, 37.87]	8.00	34.06	[31.59, 36.54]	7.83
	Sentence	35.62	[33.65, 37.60]	6.26	34.59	[32.09, 37.09]	7.91
	Read aloud	28.99	[25.55, 32.43]	10.89	30.62	[26.95, 34.28]	11.61
Paraphrase		4.07	[2.29, 5.85]	5.64	4.89	[3.31, 6.46]	5.00
Inference	Backward	5.17	[3.82, 6,53]	4.30	4.82	[3.05, 6.60]	5.62
	Predictive	1.66	[0.62, 2.70]	3.30	1.97	[1.00, 2.94]	3.07
Response	Association	2.40	[1.00, 3.79]	4.42	3.06	[1.46, 4.66]	5.07
	Evaluation	0.79	[-0.09, 1.67]	2.79	1.39	[0.41, 2.37]	3.10
	Reaction	2.18	[1.09, 3.26]	3.44	1.37	[0.49, 2.26]	2.81
Meta	Self-Monitoring	5.94	[4.48, 7.41]	4.64	6.71	[5.40, 8.03]	4.16
Comment	Text Organization	1.47	[0.46, 2.49]	3.22	1.15	[0.25, 2.04]	2.84

Note. CI = confidence interval; SD = standard deviation.

Table 4.25 shows the descriptive statistics of readers' comments based on the readers' proficiency and text condition. It is explicit in the data that Japanese EFL learners tended to engage mostly in the analyzing process during reading, regardless of the readers' proficiency. Participants produced more than 73% of the comments analyzed. Although low-proficiency readers generated very few comments on inference (high cohesion: M = 3.62., SD = 3.17, low cohesion: M = 4.35, SD = 5.36), high-proficiency readers produced more (high cohesion: M = 8.96, SD = 5.49, Low cohesion: M = 7.29, SD = 6.93).

Table 4.25

Descriptive Statistics for Process Levels of Think-Aloud Comments by Readers' Proficiency and Text Cohesion in Experiment 4

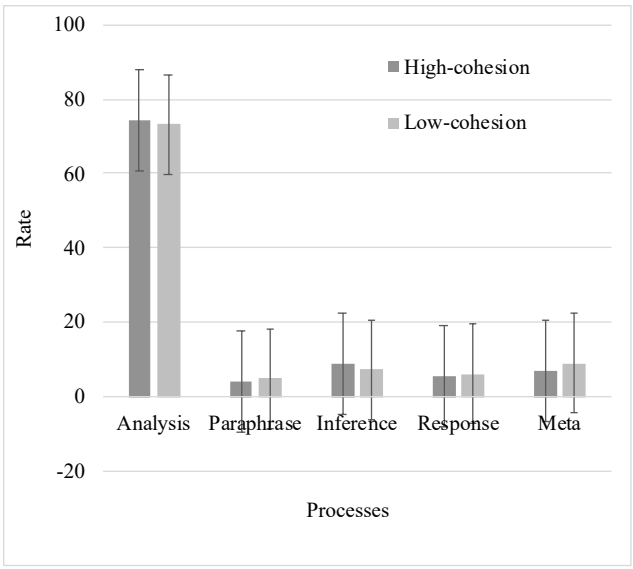
	High (n = 17)				Low (n = 23)			
	High-cohesion		Low-cohesion		High-cohesion		Low-cohesion	
Process	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Analysis	74.38	8.51	73.21	6.40	78.62	7.02	77.79	7.52
Paraphrase	4.32	6.89	4.85	5.35	3.88	4.83	4.92	4.97
Inference	8.96	5.49	7.29	6.93	3.62	3.17	4.35	5.36
Response	5.56	6.14	6.18	5.37	3.15	5.51	4.10	5.89
Meta	6.87	4.09	9.08	3.99	6.56	5.72	5.97	4.25

Note. SD = standard deviation; Meta = meta comment.

Figures 4.10 and 4.11 also describe readers' comment rates of each process by high-and low-proficiency readers, respectively.

Figure 4.10

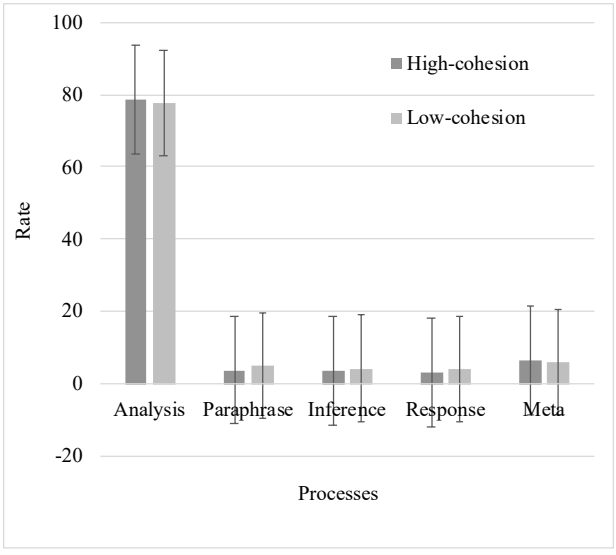
Means of Rates of Think-Aloud Comments for High- and Low-Cohesion Texts by High-Proficiency Readers in Experiment 4



Note. Error bars represent the standard errors of the means.

Figure 4.11

Means of Rates of Think-Aloud Comments for High- and Low-Cohesion Texts by Low-Proficiency Readers in Experiment 4



Note. Error bars represent the standard errors of the means.

To examine the effects of text cohesion on readers' allocation of cognitive resources, a 2 (proficiency: high proficiency, low proficiency) × 2 (text: high cohesion, low cohesion) × 5 (process: analysis, paraphrase, inference, response, and meta comment) three-way mixed ANOVA was conducted on the production rates with arcsine transformation in the think-aloud protocol task. This analysis was conducted with the reader's proficiency as a between-participant variable and text and process as within-participant variables. The result showed a significant effect of Process × Proficiency with few effect sizes, $F(4, 152) = 3.27$, $p = .013$, $\eta^2 < .01$. The results also showed that the main effect of the process was significant, $F(4, 152) = 1250.45$, $p < .001$, $\eta^2 = .96$, with a large effect size, whereas the main effect of text, $F(1, 38) = 1.12$, $p = .296$, $\eta^2 < .01$, and proficiency were not significant, $F(1, 38) = 3.10$, $p = .087$, $\eta^2 < .01$. Other interaction effects (Text × Proficiency, Text × Process, and Text × Proficiency) were insignificant as well. These results indicated that allocation of significantly different cognitive resources for different processes of reading was not affected by text cohesion and readers' proficiency.

To investigate the effect of text cohesion on readers' allocation of cognitive resources from different perspectives, a hierarchical cluster analysis was performed on the 40 participants' comments of the think-aloud protocol task. This analysis classified participants by gathering similar categories and rates of the think-aloud comments into a cluster. Readers were divided into two groups based on their reading patterns or allocation characteristics. Ward's method with the squared Euclidean distance technique was used. Through the analysis, two cluster groups, namely cluster 1 ($n = 13$) and cluster 2 ($n = 27$) were found. To check the difference in the cluster group scores of the L2 reading proficiency test, a *t*-test was conducted. The mean score of cluster 1 ($M = 16.38$, $SD =$

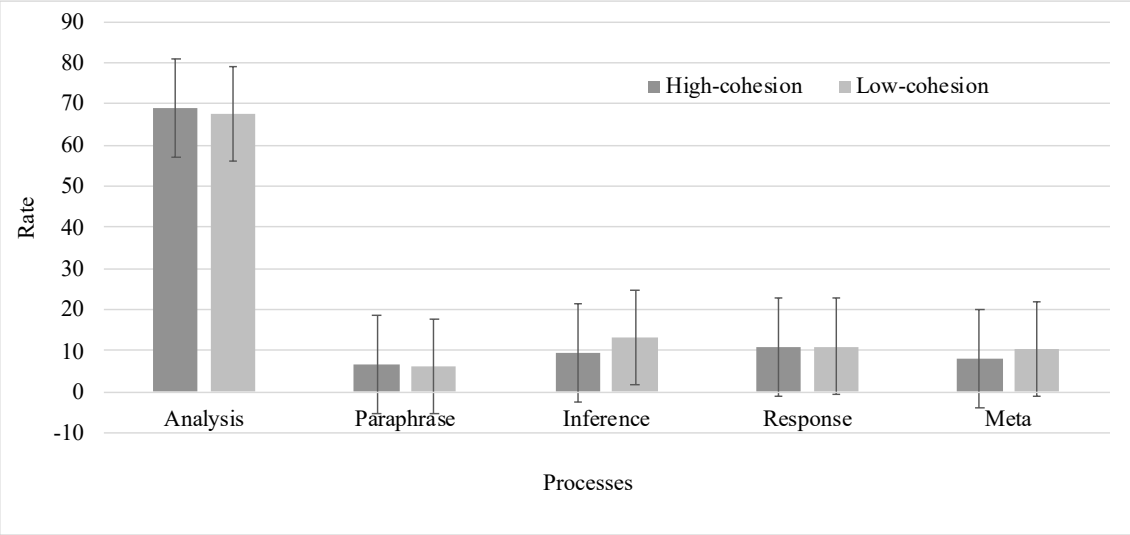
5.55, *Max* = 23, *Min* = 7) was significantly higher than that of cluster 2 (*M* = 12.51, *SD* = 4.44, *Max* = 20, *Min* = 3), $t(38) = 2.38$, $p = .023$, $d = 0.80$, and the effect size was large.

Additionally, to ensure the discriminability of the two clusters, a multivariable analysis was conducted on the 40 participants' comment rates of the think-aloud protocol task. The results showed significant differences ($p < .01$) between clusters 1 and 2 in some process-level comments (e.g., analysis, inference, association). Cluster 2 tended to be more active in the analysis process, whereas cluster 1 tended to engage more in higher processes (i.e., inference, response, and meta comment). Thus, participants in cluster 1 were labeled higher-level processing readers, and participants in cluster 2 were labeled lower-level processing readers based on their allocation of cognitive resources.

Figures 4.12, 4.13, and Table 4.26 display the production rate of the think-aloud comments per process by cluster group and reading condition (i.e., high cohesion and low cohesion). These data were also the result of the transformation of arcsine as a prior analysis.

Figure 4.12

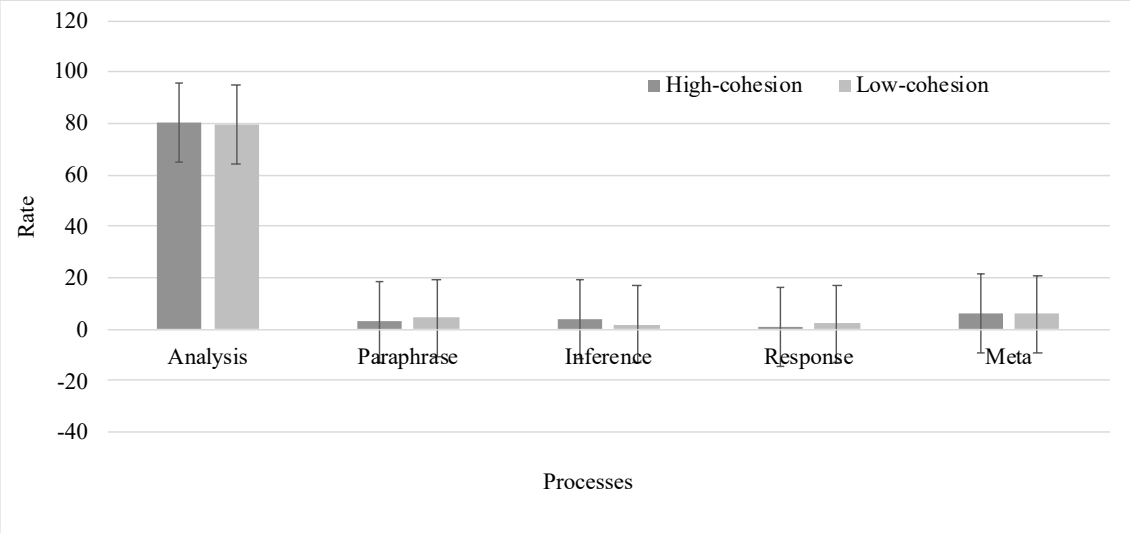
Means of Rates of Think-Aloud Comments for High- and Low-Cohesion Texts by Higher-Level Processing Readers in Experiment 4



Note. Error bars represent the standard errors of the means.

Figure 4.13

Means of Rates of Think-Aloud Comments for High- and Low-Cohesion Texts by Lower-Level Processing Readers in Experiment 4



Note. Error bars represent the standard errors of the means.

Table 4.26

Descriptive Statistics for Process Levels of Think-Aloud Comments by Cluster Group and Text Cohesion in Experiment 4

	HPG (n = 13)				LPG (n = 27)			
	High-cohesion		Low-cohesion		High-cohesion		Low-cohesion	
Process	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Analysis	69.04	5.36	67.63	4.18	80.56	5.92	79.80	4.79
Paraphrase	6.74	6.59	6.06	4.87	2.78	4.87	4.32	5.15
Inference	9.57	5.50	13.09	3.29	4.11	3.71	2.00	3.26
Response	10.94	5.11	11.09	4.14	0.91	2.25	2.04	3.63
Meta	7.98	6.28	10.33	3.48	6.07	4.31	5.83	4.03

Note. SD = standard deviation; HPG = higher-level processing group; LPG = lower-level processing group; Meta = meta comment.

As shown in Table 4.26, readers in cluster 1 (higher-level processing group: HPG) produced higher-level processing comments, such as inference, response, and meta comments. Contrastingly, readers of cluster 2 (lower-level processing group: LPG) tended to engage in lower-level processes such as analyzing words and sentences. This tendency was the same as that shown in Table 4.25 because the two cluster groups had different proficiencies, similar to the prior distinction.

To determine the effect of text cohesion on readers' allocation of cognitive resources, a 2 (cluster: higher-level processing group, lower-level processing group) × 2 (text: high-cohesion, low-cohesion) × 5 (process: analysis, paraphrase, inference, response, and meta-comment) three-way mixed ANOVA was conducted on the production rates with arcsine transformation in the think-aloud protocol task. This

analysis was conducted with cluster groups (HPG and LPG) as a between-participant variable and text and process as within-participant variables.

Table 4.27 displays the results of the ANOVA.

Table 4.27

ANOVA Results for Main Effects and Interaction Effects of Cluster and Cohesion and Process on Think-Aloud Protocol in Experiment 4

Source	*SS*	*df*	*MS*	*F*	*p*	η^2
	Between participants					
Cluster (CL)	507.87	1.00	507.87	71.54	< .001	<.01
Error	269.75	38.00	7.10			
	Within participants					
Text (T)	10.62	1.00	10.62	4.26	.046	<.01
T × CL	16.78	1.00	16.78	6.74	.013	<.01
Error (T)	94.70	38.00	2.49			
Process (P)	258171.70	2.56	101045.25	1897.61	< .001	.95
P × CL	5074.55	2.56	1986.12	37.30	< .001	.02
Error (P)	5169.93	97.09	53.25			
T × P	48.72	3.57	13.66	0.81	.507	<.01
T × P ×CL	178.92	3.57	50.17	2.98	.026	<.01
Error (T × P)	2279.36	135.53	16.82			
Total	271822.90	188.33				

The results showed the interactional effect of Text × Process × Cluster with a small effect size, $F(3.57, 152) = 50.17$, $p = .026$, $\eta^2 < .01$, and a significant interactional effect of Process × Cluster, $F(2.56, 97.09) = 1986.12$, $p < .001$, $\eta^2 = .02$, with a small effect size. The results also showed significant main effects of cohesion, process, and cluster, and a significant interaction effect of Text × Cluster (see Table 4.27). However, the interactional effect of Text × Process was not significant, $F(3.57, 152) = 13.66$, $p = .812$, $\eta^2 < .01$.

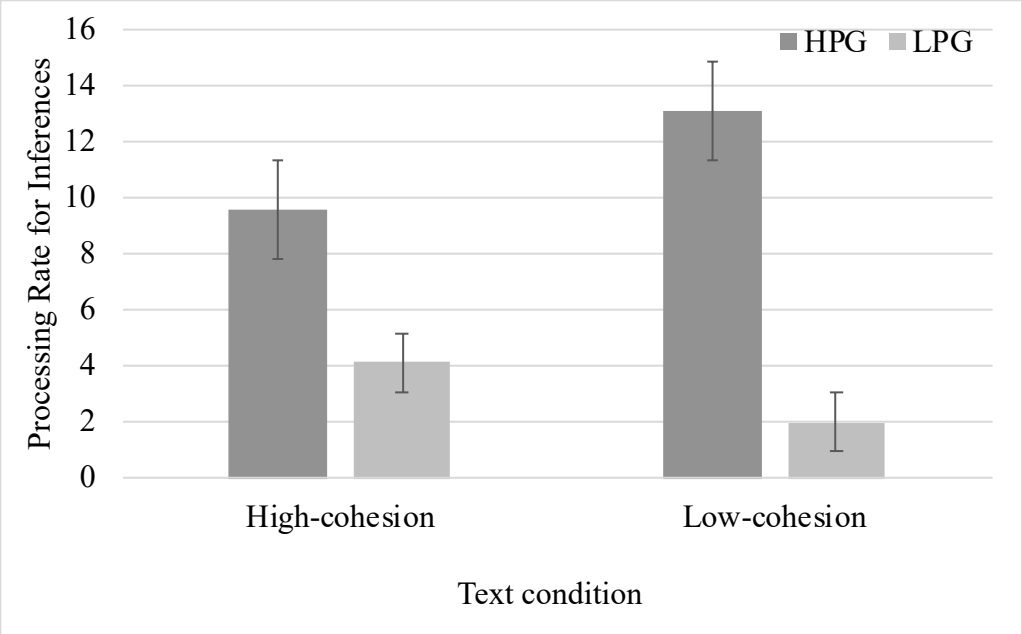
Although there were some statistically significant observations in this analysis, the main focus of the present experimental study was to investigate the effects of text cohesion.

To examine the three-way interaction of Text × Process × Cluster in more detail, the processes were further analyzed. The results showed a significant interaction effect of Cluster × Text cohesion on the inference process, $F(1, 38) = 11.59$, $p = .002$. The results also indicated that higher-level processing readers generated statistically significantly more inferences when they read the low-cohesion text ($M = 13.09$, $SD = 3.29$) compared to the high-cohesion text ($M = 9.57$, $SD = 5.50$), $p = .014$. In contrast, lower-level processing readers generated more inferences when reading the high-cohesion text ($M = 4.11$, $SD = 3.71$) than when reading the low-cohesion text ($M = 2.00$, $SD = 3.26$), $p = .031$.

These interaction results (see Figure 4.14) indicate that text cohesion affects readers' inference generation differently depending on readers' characteristics (i.e., allocation of cognitive resources).

Figure 4.14

Mean Inference Rates in Think-aloud Comments for High- and Low-Cohesion Texts by Higher- and Lower-Level Processing Readers in Experiment 4



Note. HPG = higher-level processing readers; LPG = lower-level processing readers. Error bars represent the standard errors of the means.

Further analyses by cluster also showed significant effects of Text × Process. That is, while reading high-cohesion texts, differences between HPG's comments and LPG's comments on analysis ($p < .001$), inference ($p = .001$), and response ($p < .001$) were significant. These results indicate that higher-level processing readers produced statistically significantly more comments on inference and response compared to lower-level processing readers, whereas lower-level processing readers produced more comments on analysis compared to higher-level processing readers. Similarly, while reading low-cohesion texts, readers' comments on analysis ($p < .001$), inference ($p < .001$), response ($p < .001$), and meta comments ($p = .001$) were significantly different between HPG and LPG. The higher-level processing readers generated statistically significantly

more comments on inference, response, and meta comments than the lower-level processing readers, except for one process of analysis.

Based on the above results, it was observed that text cohesion influenced readers' high-level processing (i.e., inference generation) during reading. The low-cohesion text improved higher-level processing readers' inference generation compared to the high-cohesion text, whereas the high-cohesion text facilitated lower-level processing readers' inferencing process compared to the low-cohesion text. These results indicate that readers' reading process is affected by text cohesion, as well as, readers' proficiency.

4.4.4 Discussion

Response to RQ5-1: ***Does cohesion affect inference generation among Japanese EFL readers while reading scientific expository texts?***

The result of the three-way ANOVA of readers' think-aloud protocols showed that text cohesion affects EFL learners' reading processes. In particular, higher-level processing readers, who tend to engage in higher-level processes (i.e., inference generation) during reading, generated more inferences for low-cohesion text than for high-cohesion text. In contrast, lower-level processing readers, who tend to engage in lowwer-level processes (i.e., word analysis) during reading, generated more inferences while reading high-cohesion text than low-cohesion text.

This result suggests that high-cohesion text could be facilitative for Japanese EFL readers with lower-level processing skills. In contrast, it is suggested that low-cohesion text could be beneficial for Japanese EFL readers with high-level processing skills to produce more inferences. The conducible effect of high-cohesion text on less-skilled learners observed in this experimental study corresponds with prior studies (e.g., Loxterman et al., 1994; Yano et al., 1994) and Experiment 3-B in the dissertation. In addition, the reverse cohesion effect that low-cohesion text enhances skilled readers' active inferencing process corroborated the results of previous studies that hypothesized a similar reverse cohesion effect (e.g., McNamara, 2001; McNamara et al., 1996; O'Reilly & McNamara, 2007) and Experiment 2-A in this dissertation. Below are some possible explanations for these results.

First, a possible explanation of the high-cohesion text effect on lower-level processing readers is the increased availability of information. It has been shown that readers need to infer ideas to fill in gaps in the text while reading low-cohesion texts (Hosoda, 2016; Kintsch, 1994; McNamara et al., 1996; McNamara & Kintsch., 1996;

Ozuru et al., 2010). Thus, low-proficiency readers may have difficulty in higher-level processing (e.g., inferencing). Indeed, as shown in the result section, lower-level processing readers—who had lower scores on the L2 proficiency test than higher-level processing readers—produced mostly analysis comments since they did not allocate their cognitive resources to higher-level processes. However, high-cohesion text provided readers with explicit signals and cues, which contributed to lower-processing readers' activation of identifying inference gaps within the text (e.g., Daneman & Hannon, 2001). Thus, by providing sufficient explicit relationships of ideas in high-cohesion texts, readers, who tend to engage in lower-level processes during reading and have difficulty in constructing coherent mental representations, might be able to make important inferences. This suggestion has also been proposed by Hall et al. (2015).

Second, concerning the result of the reverse cohesion effect, the same interpretation as prior studies can be regarded as reasonable. McNamara (2001), McNamara et al. (1996), and O'Reilly and McNamara (2007) suggested in their studies that high-cohesion text may not be of benefit for high-knowledge readers because such readers do not tend to use their own knowledge to integrate ideas of the text when the high-cohesion text already supplies the information needed to make inferences. In addition, high-cohesion text has the probability of making readers passive (McKoon & Ratcliff, 1992). This interpretation is supported by the results of the questionnaire in Experiment 3-A, where the data analyzed in Experiment 4 were collected. According to the descriptive statistics for the questionnaire, high-proficiency readers felt that the high-cohesion text was more difficult than the low-cohesion one, whereas low-proficiency readers felt the low-cohesion text was more difficult than the high-cohesion one. Although the interaction effect discussed here was not purely attributed to the readers' proficiency but also to their cognitive allocation grouping, it may be possible that readers with high proficiency find

highly cohesive text bothersome. This is because high-proficiency readers can infer information without elaborate revisions. Based on these explanations, it is concluded that high-proficiency readers who seem to have enough skills to generate inferences do not gain an advantage from high-cohesion texts but rather from low-cohesion texts.

In summary, based on the results of the think-aloud task, the positive effect of high-cohesion text on lower-level processing readers and the positive effect of low-cohesion text on higher-level processing readers are revealed.

Response to RQ5-2 and RQ5-3: ***Does the cohesion effect on inference generation from scientific expository texts vary as per the readers' proficiency? / Does the cohesion effect on inference generation from scientific expository texts vary depending on the readers' processing tendency during reading?***

As discussed above, the results of the ANOVA of the think-aloud task presented an interactional effect between readers and text. That is, higher-level processing readers allocate more cognitive resources in the inferencing process when reading low-cohesion text, and in turn, lower-level processing readers allocate more cognitive resources when reading high-cohesion text. From this result, one interaction effect of readers' proficiency and text cohesion was seen. This interaction effect partially corroborated previous studies (Hosoda, 2016; Linderholm et al., 2000; McNamara, 2001).

Although such an interaction was not observed for reader proficiency (RQ5-2), it was observed for readers' processing tendencies during reading (RQ5-3), which seemed to be related to reader proficiency. Since proficiency has been treated as a factor of readers in past experiments, an interaction of cohesion with proficiency was expected here as well. However, it was suggested that not simply proficiency, as measured by a reading test, but readers' processing tendencies during reading may affect the cohesion during

reading. This is one of the new findings of this study, as it has not been examined in previous studies on text cohesion.

The results are reasonable because text cohesion manipulation is related to readers' comprehension and bridging inferences during reading. Prior studies revealed that interactional effects were characterized by using difficult passages for participants (Linderholm et al., 2000), or when conducting an experimental study on low-knowledge learners (Ozuru et al., 2009). As mentioned, the experimental materials (texts) were not very difficult for the participants in Experiments 3-A and 4. Therefore, readers who were relatively proficient and could engage in higher-level processing during reading were able to actively engage in inferential activities in low-cohesion texts. Conversely, for lower-processing readers, who tended to engage in lower-level processing such as recognizing individual words and understanding their meanings, the high-cohesion texts provided markers that highlighted the relationships between individual elements and allowed readers to make inferences from elaborate pieces of information. Therefore, it is plausible that the effect of text cohesion differed in terms of the processing tendencies during reading, rather than timed reading tests, which depend on readers' knowledge of the meaning of words and speed of reading. These results indicate that text cohesion has a varying effect on drawing inferences and processing information during reading, based on the reader's processing tendency during reading. However, since this study used the think-aloud method, which relies on the spontaneous utterances of the reader, it should be noted that future research using objective measures is necessary.

4.4.5 Conclusion of Experiment 4

Experiment 4 was conducted to explore the effects of text cohesion on Japanese EFL readers' processing during reading where they were provided two versions of

cohesive texts on science topics. To extend the findings of the cohesion effects on EFL readers and overcome the limitations of past studies, Experiment 4 analyzed readers' processing while reading high-cohesion and low-cohesion science texts. By using the think-aloud method and considering readers' factors (i.e., reading proficiency and reading processing tendencies), Experiment 4 addressed the research questions below.

RQ5-1: Does cohesion affect inference generation among Japanese EFL readers while reading scientific expository texts?

RQ5-2: Does the cohesion effect on inference generation from scientific expository texts vary as per the readers' proficiency?

RQ5-3: Does the cohesion effect on inference generation from scientific expository texts vary depending on the readers' processing tendency during reading?

The results of Experiment 4 showed that text cohesion did not have a specific effect on readers' processing during scientific text reading (RQ5-1) and that reader proficiency did not alter the effect of cohesion on processing during reading (RQ5-2). Even so, depending on the reader's processing tendency during reading, text cohesion affects the reader's inference production during reading different cohesion texts (RQ5-3).

Specifically, the results revealed that higher-level processing readers generated more inferences for low-cohesion text than for high-cohesion text. In contrast, lower-level processing readers generated more inferences while reading high-cohesion text than low-cohesion text (RQ5-3). These results suggest that text cohesion has a varying effect on inference processing during reading, depending on the reader's level of processing during reading.

Based on the fact that text cohesion influences readers' processing during reading,

this study's results can be utilized for inference and strategy instruction for students facing reading difficulties. Although this study did not directly examine the relationship between text comprehension and processing of EFL readers, it would be desirable to develop future research that investigates how text coherence affects reading comprehension and the resulting learning derived from the text.

4.5 Conclusion of Study 2

Study 2 aimed to explore the effects of text cohesion (revised text in terms of cohesion) on Japanese EFL readers. It consisted of six experimental studies, including two follow-up experiments (Experiment 1; Experiments 2-A, 2-B; Experiments 3-A, 3-B; and Experiment 4). Here is a summary of the results obtained from the experiments:

Experiment 1 tested whether Japanese EFL readers were sensitive to English text cohesion, by using a cohesion judgment task. The results revealed that EFL readers were not sensitive to text cohesion when they were not provided with any specific reading direction. In Experiment 2, the effects of cohesion on Japanese EFL readers' comprehension of expository texts on social topics were investigated using a written recall task. The results of Experiments 2-A and 2-B showed that text cohesion affected readers' literal and deep comprehension when their English proficiency was at a certain level (Experiment 2-A), while low-cohesion texts facilitated readers' comprehension compared to high-cohesion texts. In addition, it was found that the effects of text cohesion on readers' comprehension differed depending on their proficiency level. When readers had adequate proficiency, low-cohesion texts improved their comprehension compared to high-cohesion texts (Experiment 2-A); however, when readers' proficiency was relatively low, text cohesion did not affect comprehension (Experiment 2-B).

Experiment 3 tested the effect of text cohesion on Japanese EFL readers using

science expository texts of different cohesions. The results suggested that text cohesion affected readers' literal comprehension when their proficiency was low, whereas there were no effects of text cohesion when readers' proficiency was high. Thus, high-cohesion text was found to facilitate comprehension of relatively low-proficiency readers (Experiment 3-B).

Experiment 4 examined the effects of text cohesion on Japanese EFL readers' processing during reading using the think-aloud method, taking into account the proficiency of the readers and their processing tendencies. The results revealed that lower-level processing Japanese EFL learners could generate more inferences in high-cohesion text than in low-cohesion text. In contrast, high-level processing readers produced more inferences in low-cohesion text compared to high-cohesion text.

The results of these experiments suggest that (a) Japanese EFL readers are not aware of the high and low levels of text cohesion unless they are given special instructions, but (b) text cohesion may significantly affect readers' reading comprehension and inference generation during reading.

Moreover, the effects of manipulating text cohesion are not necessarily the same for readers of varying proficiency levels and reading styles (e.g., high-cohesion text leads to better comprehension than does low-cohesion text). Rather, the effect of manipulating text cohesion was shown to be significantly related with the reader's proficiency, processing tendencies during reading, target text genre, and difficulty level. This dissertation, which investigated the effect of text cohesion on Japanese EFL learners' reading through multiple experiments and methods, provides valuable suggestions for EFL reading research. The findings and discussion of the text analyses and experiments are discussed in the next chapter.

Chapter 5

General Discussion

This chapter provides an overview of the findings of the two text analyses and six experimental studies to propose a comprehensive discussion of text cohesion and EFL reading. In the following three sections, I summarize the results of each study and discuss three perspectives: (a) the relationship between text difficulty and text cohesion in reading, (b) text cohesion manipulation and its effects on readers' comprehension of expository texts, and (c) readers' factors interacting with text cohesion in EFL reading and processing.

5.1 Text Difficulties and Cohesion in Reading

First, Study 1 is reviewed to discuss the relationship between text difficulties and cohesion in reading. In Study 1, the following two research questions were posed:

RQ1-1: In English language textbooks for Japanese EFL learners, how do the texts for different grade levels differ in terms of cohesion?

RQ1-2: In the EIKEN tests for Japanese EFL learners, how do the texts for different grade levels differ in terms of cohesion?

In Study 1, Text Analyses 1 and 2 were conducted to answer RQ1-1 and RQ1-2, respectively. Text Analysis 1 analyzed 222 units in 20 teaching materials and English textbooks for Japanese EFL readers (from elementary, middle, and high schools) using a Coh-Metrix text analyzer (e.g., McNamara et al., 2014). The results revealed that reference cohesion, which shows the overlap between words and concepts between adjacent sentences, in English textbooks became less cohesive as the school grade increased. However, the results for deep cohesion, which reflects how well the

information, ideas, and events in the text are related by connectives, were not linear; high school textbooks, which increased in text length, had the highest deep cohesion.

Expressly, the answer to RQ1-1, specifically, the reference cohesion, shows that the more difficult the English sentence, the lower the cohesion, and the text is less explicit for the reader. However, deep cohesion, which is reflected by the amount of causal and intentional conjunctions, was the most cohesive in high school textbooks with a large amount of vocabulary and content.

Further, Text Analysis 2 analyzed reading passages in the EIKEN test (Grades 1–3), a standardized English test common in Japan, to investigate texts other than English textbooks. The results of analyzing 501 texts in EIKEN tests by Coh-Metrix suggested that reference cohesion decreases as the text becomes more difficult. Similarly, deep cohesion tended to decrease as the text became more difficult (from Grade Pre-2 to Grade 1). However, it was confirmed that cohesion was low even in Grade 3, the easiest grade of the texts analyzed.

Regarding the reference cohesion in RQ1-2, which is the degree of overlap of content words and the same concepts in a text, the degree of overlap decreases as the English text becomes more difficult and less cohesive. In contrast, deep cohesion, which is reflected in the amount of conjunctions in the text, did not change linearly with the difficulty of the text. Regarding the EIKEN texts, the deep cohesion was relatively high for the texts in higher grades, while it was low for those in intermediate grades.

Considering the results of Text Analyses 1 and 2, the same tendency was confirmed for reference cohesion. That is, the higher the difficulty and level of the text, the lower the reference cohesion, and the lower the difficulty, the higher the reference cohesion.

This is reasonable given that easier texts use simple sentences that repeat certain words and phrases for younger and less proficient readers. However, in more difficult

texts, paraphrasing, pronominalization, and ellipsis are more likely to occur than the simple repetition of nouns, and the reference cohesion is likely to be lower. These results are consistent with previous research on L1 (e.g., McNamara et al., 1996), which argued that readers with poor knowledge often struggle to understand difficult texts because the elements described are not explicit or lack relational connectives; such low-cohesion texts are considered to be difficult for readers. This is consistent with claims that text cohesion is a significant aspect of text difficulty (McNamara et al., 2012; McNamara et al., 2014).

However, for deep cohesion, which is determined by the use of connectives of causal and intentional relationships, no consistent results were observed in Text Analyses 1 and 2. This could be due to the differences in the materials used in the analyses. Text Analysis 1 analyzed textbook texts including dialogue. Some textbooks were organized using dialogue-based texts for learners, and I included such texts for grasping tendencies in English textbooks. In contrast, Text Analysis 2 included more explanatory texts than the textbooks used in Text Analysis 1. Because Text Analysis 2 was conducted using EIKEN test passages, the analyzed materials were varied from short passages to long difficult expository texts from different grade levels.

Taken together, it is suggested that the reference cohesion can be a predictor of EFL text readability (i.e., the higher the reference cohesion, the easier it is for EFL readers), but the deep cohesion can interact complexly with difficulty and other text features. Indeed, the results for deep cohesion were not consistent between Text Analyses 1 and 2. This implies that the relation between deep cohesion and text difficulty was not as simple as that of reference cohesion without considering text genres, which determines the use of connective expressions in texts. This is reasonable considering that the different genres of texts for L1 readers differ in cohesion indexes (Lightman et al., 2007). Deep cohesion is sensitive to the use of connective expressions in a text, and text genre affects the writing

style of a text. Thus, Text Analyses 1 and 2, which focused on different genres, yielded inconsistent results. I have to admit the limitation of Study 1 was that the analysis did not consider text genre or topics. Therefore, further studies should be conducted on this matter. Deep cohesion and readability relations need to be discussed in text analyses controlling for text genre factors.

While I cannot conclude the relationship between deep cohesion and text readability for EFL materials, reference cohesion can be a possible predictor of readability. The present finding about reference cohesion and text difficulty is in accordance with Nahatame (2021). He examined the relationship between cohesion and difficulty of reading materials for young L2 learners, and revealed that multiple indices of cohesion contributed about 40% to the prediction of text difficulty but cautioned that cohesion did not completely explain it. It cannot be concluded that text readability is completely attributable to the reference cohesion of a text; however, we can say that the reference cohesion index is an essential index for estimating reading difficulty for L2/EFL readers.

However, in L1 and ESL research, reverse tendencies were observed. That is, the reference cohesion increases across grade levels (McNamara et al., 2012; Plakans & Zeynep, 2016). Thus, it should be noted that the present findings can be applied to EFL reading materials in which some manipulations or word/grammar controls are conducted compared to those of ESL or L1 readers. Thus, further large-scale studies considering material language (as L1/ESL/EFL) and grade levels (e.g., college-level) need to be conducted.

5.2 Manipulation of Text Cohesion and Readers' Understanding

Based on Study 1, Study 2 examined the effect of the strength of text cohesion on reading comprehension. The study was composed of six experimental studies. By conducting several experiments, Study 2 answered 11 research questions. To discuss the effects of cohesion manipulation on L2 reading, I will summarize the results of each study.

First, Experiment 1 set the following research questions:

RQ2-1: Are EFL readers aware of the degree of cohesiveness of a text?

RQ2-2: Do EFL readers judge high- and low-cohesion texts differently in terms of their proficiency level?

The cohesion judgment task was conducted in Experiment 1 to test readers' sensitivity to cohesion. Two cohesion versions of texts were used for expository texts of scientific content. One was a low-cohesion text with less use of connectives and a lesser degree of noun redundancy; the other was a high-cohesion text with a high use of connectives, noun repetition, and additional information. University students of Japanese EFL learners read high- and low-cohesion texts and judged their cohesiveness. The results suggested that although readers gave a higher cohesive rate to high-cohesion texts and a lower rate to low-cohesive texts, the difference in ratings was not statistically significant (RQ1-1). For RQ1-2, the results also showed that there was no effect of reader proficiency on cohesive judgment.

The results are reasonable considering that Japanese EFL readers may not be sensitive to the English writing style and may not make good use of linguistic cues in texts (Ozono & Ito, 2003). Participants were not given any pre-reading instructions to grade cohesion before reading the texts. Therefore, the readers were possibly focused on text comprehension and did not set their reading criteria on cohesion, which is a textual

feature.

L1 and L2 reading studies dealing with reading instructions that give readers a specific reading goal or direction suggest that such instructions affect readers' standards of coherence (e.g., van den Broek et al., 2011). A standard of coherence is defined as "the degree of comprehension that a reader attempts to achieve during the reading of a text" (Linderholm et al., 2004, p. 168), and affects readers' processing and comprehension (e.g., Ushiro et al., 2018). Indeed, Nahatame (2017), who investigated Japanese EFL readers' employment of standards of coherence, suggests that EFL readers can be sensitive to the coherence of intentional and causal relations between sentences. Given that participants in Experiment 1 received instruction to read experimental texts for their understanding, the result that they could not significantly evaluate different cohesiveness was reasonable.

Nonetheless, text cohesion is a necessary element for readers to construct a coherent understanding of the text (Halliday & Hasan, 1976; McNamara et al., 2014). Although there have been L1 and L2/EFL reading studies that have investigated the effects of cohesion manipulation and strategy instruction on reading (e.g., Kitajima, 1997; McNamara et al., 1996; Horiba, 2017; Hosoda, 2016; Safaie, 2020), there is still room for an investigation of readers' sensitivity to text cohesion. Thus, to expand and validate the present results, future studies concerning readers' awareness of meta-linguistic knowledge and cohesive devices are needed. Additionally, text cohesion is the degree of cohesiveness of texts, and readers' awareness and judgments can vary. Thus, various types of cohesive ties (e.g., additive, temporal, and causal connectives, van Silfhout et al., 2015) and text genres should be used in further studies.

Experiment 1 showed that the degree of text cohesion was not statistically significant in situations in which the reader was not instructed to pay special attention to it. However, regardless of the reader's awareness of the meta-features, structure and

vocabulary can influence the reader's understanding. In Experiments 2 and 3, the effects of text cohesion manipulation on readers' comprehension were investigated. Furthermore, by using non-scientific and scientific topics as materials for each experiment, the effects of different text genres were tested. Moreover, the researcher investigated the effects of readers' reading proficiency levels. Experiments 2 and 3 answered the following questions:

RQ3-1: Does high or low cohesion affect Japanese EFL readers' surface-level comprehension of social expository texts?

RQ3-2: Does high or low cohesion affect Japanese EFL readers' deep-level comprehension of social expository texts?

RQ3-3: Does the cohesion effect on the comprehension of social expository texts vary as per the readers' proficiency level?

RQ4-1: Does high or low cohesion affect Japanese EFL readers' surface-level comprehension of scientific expository texts?

RQ4-2: Does high or low cohesion affect Japanese EFL readers' deep-level comprehension of scientific expository texts?

RQ4-3: Does the cohesion effect on the comprehension of scientific expository texts vary as per the readers' proficiency level?

To answer RQs 3-1, 3-2, and 3-3, high-cohesion and low-cohesion versions of non-scientific texts were used in Experiment 2. Japanese EFL readers read two texts of either version and completed comprehension measurement tasks, such as a written recall task. In Experiment 2-A, which was conducted on a relatively high-proficiency group, the results showed that the low-cohesion text facilitated readers' comprehension more than

the high-cohesion text. However, Experiment 2-B, which was followed by students from another university with a lower relative reading proficiency than Experiment 2-A, did not confirm any difference in comprehension.

These results indicate that the effect of cohesion on the comprehension of similar materials differs depending on the reader's proficiency. This can be discussed from the perspective of the reverse cohesion effect, which has been confirmed in previous studies on L1 (McNamara, 2001; McNamara et al., 1996; McNamara et al., 2011; O'Reilly & McNamara, 2007; Smith et al., 2021). Generally, increasing text cohesion can promote readers' comprehension (McNamara et al., 2012). The materials used in Experiment 2 described campaigns and issues related to nature and education and did not contain much technical content. For such cases, it was found that low-cohesion texts can promote comprehension rather than increase cohesion for readers with a certain level of proficiency. However, for the less proficient readers who participated in Experiment 2-A, no significant difference in comprehension by cohesion was confirmed. This study did not compare different proficient readers in the same experiment and cannot strongly conclude, though can suggest, that the effect of high-cohesion texts on comprehension may be observed when targeting less proficient readers. Although Experiments 2-A and 2-B targeted only a limited number of readers of various proficiency levels, the findings are valuable as they reveal that the effect of cohesion on text comprehension differs depending on the reader's proficiency level in English.

Further, Experiment 3 examined the effect of cohesion on EFL readers' comprehension of expository texts of scientific explanations. As in Experiment 2, to answer the research questions, Experiment 3-A targeted relatively proficient undergraduate and graduate students, and Experiment 3-B targeted relatively less proficient undergraduate Japanese EFL learners. In Experiment 3-A, there was no

significant difference between the high-cohesion and low-cohesion texts in comprehension, as measured by the readers' written recall task. Expressly, there was no specific effect of either cohesion text to promote better comprehension. However, Experiment 3-B, a follow-up experiment with less proficient learners, suggested that less proficient readers specifically achieved better comprehension in high-cohesion science texts.

These results may seem surprising and different from those of Experiment 2, but they are reasonable as the experimental material was difficult (science texts used in Experiment 3). Scientific texts are more difficult to read than general texts because they present new concepts or explain specialized content or complex scientific phenomena. Therefore, it is assumed that modifications that enhance cohesion are more likely to promote readers' comprehension. Since Experiments 3-A and 3-B targeted many students from fields other than science, it is likely that the cohesion-enhancing revision would have contributed more to comprehension. Furthermore, as indicated by Experiment 2, the effects of cohesion varied depending on the reader's proficiency level. Thus, it is likely that high-cohesion texts promoted comprehension when the proficiency level was low, and the effect of cohesion was not as great for those with some proficiency. However, based on previous studies, when the readers were more proficient or specialized than the participants in Experiment 3-A, their comprehension may be higher in low-cohesion texts than in high-cohesion texts (i.e., observation of direct reverse cohesion effects).

To summarize, it is possible that in a reading environment with no specific instructions, readers do not pay attention to the meta-information (or text cohesiveness, which is how explicitly the text presents the information) in the text (Experiment 1). Text cohesion affects readers' comprehension regardless of whether they perceive cohesiveness (Experiments 2-A, 2-B, 3-A, 3-B). The results suggest that low-cohesion

texts may facilitate comprehension when the reader's proficiency level is high, while high-cohesion texts may enable comprehension when the reader's proficiency level is low. Moreover, the topic (familiarity with the reader) may also affect the cohesion effect; if it matches the reader at a suitable level of difficulty, comprehension may not be affected by the difference between high -and low-cohesion texts.

The results of experimental studies suggest that Japanese EFL reading and comprehension can be affected by text features. These results are supported by reading theories (e.g., Kintsch, 1988; van den Broek et al., 1999). Theoretically, readers are assumed to generate inferences to construct coherent mental representations (e.g., Graesser et al., 1994; van Dijk and Kintsch, 1983). In addition, readers generate bridging inferences to fill in referential or situational cohesion gaps (McNamara & Magliano, 2009). The landscape model assumes readers' strategic and deliberate retrieval during reading (coherence-based retrieval, Linderholm et al., 2004). Considering that text cohesion is expressed by explicit cues such as connectives and argument overlaps, readers' unconscious or conscious processing of information mapping or retrieval can be affected by the degree of text cohesion. Thus, it is reasonable to assume that manipulating text cohesion affects readers' comprehension and processing (e.g., Hosoda, 2016; McNamara & Kintsch, 1996; van Silfhout et al., 2015).

These theory-based assumptions have been investigated mostly in L1 reading research, and there have been limited studies on the effects of text cohesion manipulation on L2/EFL reading comprehensively. This study extended the previous L1 findings, that is, positive or negative (reverse) cohesion effects on text comprehension (e.g., Gilabert et al., 2005; Linderholm et al., 2000; Loxterman et al., 1994; Ozuru et al., 2010), in the EFL reading context. Additionally, considering text genre and readers' proficiency, this thesis highlighted interesting mutual effects of text cohesion, text-related (genre), and reader

factors (proficiency and processing tendency). Expressly, text cohesion affects EFL readers' conceptual recall and inference generation during reading, and may be related to the topic of the text, readers' background knowledge, and proficiency level. Although this may have been observed because the study controlled for the difficulty of the vocabulary in the texts, it is an interesting finding that cohesion can promote or hinder comprehension in different combinations of genres and proficiency levels.

This study emphasizes the importance of text cohesion in EFL reading. However, it should be noted that cohesion manipulation conducted in experiments had several perspectives (i.e., argument overlap, connectives, explicit explanation), and this study could not conduct completely controlled cohesion manipulation in terms of the lexical, grammar, and organization of experimental texts. Given that text cohesion is represented by grammar, lexicon, and their connective expressions (Halliday & Hasan, 1976), future studies need to investigate the effects of manipulating each linguistic element that forms text cohesiveness. In addition, although the present study conducted multiple experiments with readers of different proficiency levels, it should preferably be tested with a large-scale experiment. Therefore, although the findings of this study are valuable in suggesting the importance of cohesion in EFL reading, it should be noted that there are some limitations that need to be addressed in future research (see next chapter).

5.3 Readers' Factors in L2 Reading Processing and Comprehension

Finally, the relationship between reader factors and L2 reading is discussed in this section. In Experiment 4, the researcher tested whether different text cohesion changes the way readers process the text during reading.

RQ5-1: Does cohesion affect inference generation among Japanese EFL readers while reading scientific expository texts?

RQ5-2: Does the cohesion effect on inference generation from scientific expository texts vary as per the readers' proficiency?

RQ5-3: Does the cohesion effect on inference generation from scientific expository texts vary depending on the readers' processing tendency during reading?

To answer the RQs, a think-aloud method was used to observe readers' processing during reading. In this task, the participants were asked to speak what they were thinking or what came to mind while reading the text. As the task load during reading was relatively high, the experiment was conducted on relatively proficient university and graduate students, as in Experiments 2-A and 3-A. Scientific text was also used. The results showed that readers' processing during reading (inference generation) was affected by their interaction effect and cohesion. Specifically, those who tended to perform more higher-level processes during reading generated more inferences for coherent reading when reading low-cohesion texts compared to high-cohesion texts. This is consistent with the L1 study's claim that low-cohesion texts promote spontaneous inferences if the reader is skilled. However, for readers engaged in lower-level processes during reading, such as word decoding, it was found that high-cohesion texts were more supportive of inference

generation.

Notably, the effect of different cohesiveness on reading processing was confirmed in Experiment 4, in which the participants were divided based on their processing tendencies (higher-level processing readers, lower-level processing readers) and level of proficiency during reading. While Experiments 2 and 3 suggested interactional effects between readers' proficiency, cohesion, and text genre, Experiment 4 indicated that there was an interactional effect of cohesion and reader processing tendencies (which had a certain relationship with proficiency) on inference generation during reading. Although Experiment 4 did not directly explore the relationship between processing and text comprehension in different cohesive texts, readers' inference generation can change when they are given texts with different cohesiveness, which is a new finding in EFL reading research.

Previous reading research has suggested that proficiency and the reader's ability to make inferences are critical for successful reading comprehension (e.g., Graesser et al., 1994; McMaster et al., 2012; Yuill & Oakhill, 1991). These skills are considered higher-level processing and are characterized by proficient readers in both L1 and L2 (e.g., Horiba, 1996, 2000; McNamara & Kintsch, 1996). However, L2/EFL readers lack lower-level skills and linguistic knowledge. For example, Jeon and Yamashita (2014) pointed out that L2 grammar, vocabulary knowledge, and decoding skills play important roles in L2 reading comprehension.

The present study demonstrated that readers' benefits received from different texts differ according to their reading proficiency. Moreover, it was found that the tendency of readers to allocate their cognitive resources in reading was related to their L2 proficiency and played different roles in processing different cohesive texts.

As mentioned above, these two factors can be related: readers who engage in

higher-level processing (inferencing) are more proficient than those who engage in lower-level processing (decoding). It is very difficult to determine whether the reader's skill comes first, or whether the reader's strategic processing comes first. However, the fact that these profiles of EFL readers affect their processing and comprehension of texts is one of the interesting findings of this study. This result can be explained by the readers' standards of coherence (e.g., Linderholm et al., 2004; van den Broek et al., 1995). The strength and type of standard of coherence, which is the degree of comprehension that a reader tries to achieve in reading (e.g., Linderholm & van den Broek, 2002; van den Broek et al., 2001), can vary depending on the reader's L2 reading proficiency (e.g., Nahatame, 2017). The present study supports this finding, given that readers' reading proficiency and their processing tendencies affect the positive/negative effects of increasing text cohesion.

Unfortunately, as the study concerned the limited material and proficiency of readers, it was not possible to determine the threshold for the effect of high-cohesion or low-cohesion texts on readers' comprehension and processing. Moreover, it was not possible to consider and analyze all the factors in one empirical study, although they were comprehensively examined. However, the study revealed that text coherence may be an important factor related to difficulty not only in L1 but also in EFL, and that it interacts with other text factors (genre) and reader factors (proficiency, processing, and distribution during reading) affecting text comprehension and processing. These findings show that text cohesion affects Japanese EFL learners' processes or comprehension of reading in a complex combination of factors.

Chapter 6

Conclusion

6.1 Summary of Findings

This dissertation consists of two main studies (Study 1 and Study 2) that explore text cohesion in teaching materials and textbooks for Japanese EFL readers and its effects on their cohesion judgment, comprehension, and processing. Two text analyses were conducted under Study 1 (Text Analyses 1 and 2), and four experiments and two follow-up experiments were conducted under Study 2 (Experiments 1–4).

Text Analyses 1 and 2 examined the cohesion of textbooks designed for Japanese learners of English from elementary to high school and the texts of large-scale examinations widely used by them. The results suggest that reference cohesion is related to textual difficulty. That is, the more difficult the reading material is, the lower reference students' cohesion tends to be.

Experiment 1 investigated whether Japanese EFL readers were sensitive to English text cohesion using a cohesion judgment task, the results of which suggested that EFL readers are not sensitive to the cohesiveness of the text when they are not provided any specific reading direction. Experiments 2-A, 2-B, 3-A, and 3-B examined whether text manipulation (high-/low-cohesion texts) affected readers' comprehension of the historical-topic and scientific-topic expository texts. Results imply that low-cohesion texts facilitated readers comprehension when readers' proficiency was high or the text topic was familiar to readers; high-cohesion texts were beneficial for low-proficiency readers and when reading unfamiliar texts like science texts. Experiment 4 examined readers' processing when reading different cohesive texts, implying that text cohesion facilitated readers' inference generation differently depending on their cognitive-

allocation tendencies during reading.

In conclusion, this dissertation revealed that text cohesion can be one of the difficulty indexes in EFL reading materials and that it affects EFL reading processing and comprehension both positively and negatively depending on readers' proficiency, cognitive-allocation tendencies, and text genre.

6.2 Limitations of the Present Research and Suggestions for Future Research

Although the present study may present new findings in EFL reading research, it faces some limitations. These include: (a) the methodologies used, (b) the sample sizes and variety of material used, and (c) participants' characteristics and individual factors.

The first limitation relates to the methodology used in the present study. While this study used appropriate measurement methods for each experiment, future studies could improve on these. I used written recall and why-question tasks (Experiments 2–3 in Study 2) to measure readers' text comprehension. As written recall tasks do not assess readers' situation model comprehension (e.g., Koda, 2005), a why-question task was used to measure readers' ability to construct situation models. However, as the questions were designed by the researcher alone, the validity of the task could be doubtful. Some answers to the why-question task did not rely on the textual content but rather on the readers' knowledge. Such answers were considered bad ones and were rated with low scores because they did not answer based on the text content despite the task instruction. However, more appropriate tasks to measure readers' deep understanding of texts are desirable. To assess readers' comprehension outcomes from reading more accurately, well-designed comprehension tasks with high validity and reliability should be used in future studies.

Moreover, this study adopted a think-aloud task to assess readers' online processing

during reading (Experiment 4 in Study 2). This is an established method in reading research to observe readers' spontaneous processes and their standards of coherence (e.g., comprehension monitoring, inference generation) during reading (e.g., Horiba, 2013, Ushiro et al., 2018). However, this method analyzes only the processing reported/expressed verbally by readers; we cannot deny the possibility that there is processing that they did not externalize. Because the think-aloud protocols reflect parts of readers' attention and verbalized processing during reading, we should be careful in interpreting the results of the think-aloud task. Further, although it is reported by Leow and Morgan-Short (2004) that the existence of the think-aloud task (reading with/without the task) does not affect the degree of L2 readers' comprehension outcomes, researchers should be aware of its introspective and concurrent task features, especially when targeting nonproficient readers who have little space to conduct secondary tasks simultaneously during reading. Additionally, future study needs to consider task instruction effects on readers' meta cognitive processing during reading.

Initially, this study aimed to observe the online processing of readers through experiments using reading time and eye-tracking measurement, but experiments could not be conducted to collect such ecological data due to the explosive infection of COVID-19 and concerns pertaining to infection prevention. In the future, where the situation permits, the readers' processing during reading and their responses to different cohesive texts should be monitored by combining objective and ecological data such as reading time and eye movements.

The second limitation pertains to the material used in the present study. In Study 1, textbooks and EIKEN reading texts were analyzed. However, the analyses conducted on texts including different structures and topics (i.e., dialogue, historical, descriptive) and this study could not account for such differences. I have to admit that text analyses for

cohesion and readability in terms of different topics, genres, and structures would be desirable and my study has limited variation in this regard.

Additionally, Study 2 dealt with two types of materials: scientific texts (explaining how heat conduction and photosynthesis work) and general explanatory texts (describing a bear eating garbage or a project to save poor children). Previous studies have rarely examined the effects of cohesion, including text genre, and this study provides suggestions for such studies. However, it is difficult to generalize the results to other texts as there is still room to manipulate the number of words, difficulty level, topic, and genre of the texts. Specifically, in L1 studies, reading scientific texts is considered difficult and text cohesion has been manipulated to help readers' understanding (e.g., McNamara et al., 1996, 2011). However, for L2 and EFL readers who lack knowledge and experience, cohesion manipulation may be useful for narrative texts, which are generally regarded as gentler than expository texts in reading comprehension studies, as well as for historical texts. Future research should examine this with a wider range of text genre and difficulty level texts. Considering that reading comprehension consists of the interaction between the reader and the text, how readers respond to any text (and its cohesion manipulation) should be examined.

Third, further studies should consider participants' individual differences (e.g., proficiency and reading style) more carefully through large sample experiments (Hyönä & Nurminen, 2006). Although the present study aimed to test the effects of cohesion on Japanese EFL readers with different English proficiency levels (Experiments 2-A, 2-B, 3-A, and 3-B), the participants were different in each experiment, and their English proficiency could not be measured by the same standardized reading comprehension test (different reading proficiency tests or their self-reported test scores were used because participants' general proficiency differed). Therefore, to verify readers' proficiency

factors more accurately, it is necessary to collect learners with a range of proficiency levels and verify the effects of cohesion in their reading in larger-scale experiments. Moreover, if the experimental environment is permitted, other reader factors that may affect reading comprehension should be examined. For example, the reader's working memory capacity, which may affect processing during reading, and reading style, such as the kind of information the reader pays attention to and how they allocate their attention, should also be considered. Future studies should test the effects of text cohesion on a greater range of Japanese EFL learners whose proficiency and other individual factors differ more broadly.

6.3 Theoretical Implications

Although the present study faces some limitations, it also contributes some theoretical and pedagogical implications. The theoretical implications of the study can be categorized as follows: (i) implications regarding text cohesion and the readability of texts for EFL readers, and (ii) implications regarding EFL reading processes and comprehension of different cohesive texts: the interaction of text cohesion, genre, and reader factors in EFL reading.

6.3.1 Text Cohesion and the Readability of Texts for EFL Readers

The results reveal the characteristics of cohesion in English materials for EFL readers: the more difficult the text is, the lower the reference cohesion tends to be, and the easier it is for the text to have a higher reference cohesion. Prior research has proposed the importance of considering text cohesion for readers and the idea that it is an important readability indicator in native reading (McNamara et al., 2012). Although limited studies have investigated text cohesion in EFL contexts, the present study confirms that reference

cohesion seems to be associated with difficulty and could be a readability factor in EFL materials. This study will be a valuable first step in EFL reading research. Although their study analyzed English language materials for high school and college students, Nahatame (2021) recently pointed out that text cohesion can be related to the difficulty of materials for younger students. The development of AI, automatic English evaluation, and related research is advancing every day. More research is needed to explore details of the relationship between text cohesion and readability in L1 and L2/EFL English texts.

Moreover, text cohesion and readability should be examined not only in reading comprehension research but also in other research fields. For example, in testing, more discriminative tests can be created by considering text cohesion in passages that measure inferential ability in reading. Although studies have analyzed the semantic relatedness of words in vocabulary and reading comprehension tests (e.g., Hamada, 2013), the applicability of text cohesion editing to the measurement of readers' inferential ability remains to be clarified. Furthermore, regarding difficulty adjustment for readers, text cohesion can be related to difficulty in adaptive testing. It can also be applied to the analysis of learners' English written essays by teachers and learners for reader-conscious writing instruction. It will be necessary to consider text cohesion as an important difficulty-related indicator not only in reading comprehension research but also in research in related fields.

6.3.2 EFL Reading Processes and Comprehension in Different Cohesive Texts: Interaction of Text Cohesion, Genre, and Reader's Factors in EFL Reading

The present study also empirically demonstrated that EFL reading comprehension is affected by text cohesion. Specifically, in scientific text reading, high-cohesion texts may improve comprehension outcomes among less proficient readers. Low-cohesion

texts contribute to better comprehension when a reader of relatively high proficiency reads topics that are more familiar to the reader. Our results also indicate that the degree of text cohesion can affect the inference production of readers with different processing tendencies during reading comprehension.

Reading comprehension is achieved through an interaction between the text and the reader, and the present study revealed that textual factors affect learners' comprehension and processing of texts in EFL reading. Furthermore, reading comprehension is also achieved by manipulating the degree of cohesion on EFL readers' proficiency and the tendency to distribute their cognitive capacity to processes while reading.

If the goal of reading comprehension is to build rich situation models based on coherent text comprehension, then it is important to find textual factors that readers cannot change regarding their influence on text comprehension and the generation of inferences in EFL reading. These discussions have been active in L1 reading research, but it is rare to find research that takes into account material topics and reader factors in EFL reading.

For EFL reading, it has been shown that readers tend to devote more cognitive resources to lower-level processing while reading than native speakers (Horiba, 1996), and that the processing and success of inference generation differ depending on the proficiency level (Shimizu, 2009; Hosoda, 2014; Nahatame, 2014). However, only a limited number of studies have comprehensively examined the interaction between texts and readers. Further studies are needed to examine the reverse cohesion effects in EFL reading comprehension in more detail (i.e., the difficulty level, topic, and reader). Along with texts and readers, some reading models suggest that task effects should be considered in reading (e.g., Rouet & Britt, 2011). Whether text cohesion contributes to EFL readers'

better comprehension in combination with reading situations and purposes needs to be examined.

Although there are some issues that need to be examined in future studies, the possibility of text cohesion affecting the comprehension and processing of EFL reading indicates that future reading studies should consider it as a textual factor that may affect the performance and outcome of reading. In future EFL reading studies, text cohesion should be considered a textual factor in the adjustment of material difficulty and control.

6.4 Pedagogical Implications

The pedagogical implications of this dissertation are summarized as follows: (i) teaching text cohesion for better comprehension in EFL reading, (ii) teaching inference generation in reading for EFL readers, and (iii) selecting or developing suitable materials regarding cohesion for EFL readers.

6.4.1 Teaching Text Cohesion for Better Comprehension in EFL Reading

This study indicates that Japanese EFL readers may not be sensitive to the connections and explicitness of words and sentences in a text (i.e., text cohesion) if they are not instructed to pay attention to textual features. Even if it is implicit to the reader, it can still affect their processing and understanding. If they are consciously aware of text cohesion and can spontaneously engage in reasoning to link information together when they feel a lack of connection, they may be able to achieve reading comprehension more efficiently. Hence, reading comprehension instruction should focus on text structure (e.g., title and supplementary information) and discourse markers (e.g., *for example*), and use them as clues to help students relate and understand elements in the text.

According to Kitajima (1997), L2 learners who are instructed to focus on referential

relations in text comprehend text better compared to those who do not receive this instruction. Thus, it is important to raise students' awareness of linguistic cues in texts and expand their strategy use during reading. Although Kitajima's study pertains to learners of Japanese, Ozono and Ito (2003), along with other similar studies (e.g., Aidinlou et al., 2012; Safaie, 2020), find that Japanese learners of English may not make good use of linguistic cues in texts. Given this, it could be effective to have learners focus on the text structure (e.g., the title and mental information) and discourse markers, and train them to relate to and understand the elements using these cues.

For instance, to make learners aware of the relationship between sentences in a text, teachers can ask them to consciously pay attention to the connectives and discourse markers and have them think about how they connect the elements of the text (e.g., group work). Alternatively, teachers can leave the discourse markers in the text blank and ask learners to fill in the blanks.

In this way, learners can experience how elements in a text are interconnected. We must be aware that text cohesion is not a solution to all the problems seen in EFL reading (Carrell, 1982). However, it may contribute to broadening the reading comprehension of EFL readers to allow students to experience how text cohesion, which makes it coherent, is represented.

6.4.2 Teaching Inference Generation in Reading for EFL Readers

It is important to improve learners' reading skills to adapt to various texts (Beck et al., 1991; Graesser et al., 2003). Indeed, Grabe and Stroller (2019) point out that reading curriculums should be "(a) helping students expand their vocabulary and develop independent vocabulary-learning strategies, (b) building students' discourse structure awareness to support comprehension, (c) providing students with opportunities for

reading-fluency improvement, and (d) training students to become strategic readers while they are reading for comprehension" (p. 143).

Thus, educators must provide language learners with adequate training to use effective reading strategies (e.g., comprehension monitoring, self-explanation reading training, McNamara, 2017) as the skill of using effective strategies is key to successful reading (Ozuru et al., 2009). Nevertheless, as O'Reilly and McNamara (2007) point out, there has been a lack of instruction regarding both teachers' strategies and students' engagement in various strategic processes.

In relation to this study, the results suggest that EFL college readers can infer across gaps even when text cohesion is low. Specifically, readers with enough proficiency can overcome low-cohesion texts and generate many inferences, while readers with low proficiency and limited capacity for higher processing make use of inferencing in the explicit high-cohesion texts. Although low-cohesion text is challenging for poor readers, it is suitable for proficient readers, who fill in the gaps in the text. This is because low-cohesion texts leave room to infer gaps and facilitate readers' inference generation (e.g., Dahl et al., 2020; Radvansky et al., 2014), and readers with a certain proficiency can engage in higher-level processing during reading (e.g., Horiba, 2000). Thus, low-cohesion texts can be used to foster readers to be *autonomous readers*, who can read on their own and interact with a text. By providing such readers with low-cohesion texts, they can try to become more responsible and strategic readers.

Considering that, it may be necessary to improve readers' reasoning ability by giving them cohesive manipulated texts to train them to read between the lines or to read materials with gradually decreasing cohesiveness (e.g., deleting discourse markers in a text and making readers infer the relations between sentences). Although the relationship between inference instruction and use of different cohesive texts in EFL reading needs

further research, it should be considered in terms of instruction.

Note, however, that the ultimate goal of reading instruction is to raise learners to be autonomous readers who can read authentic texts on their own, even if these texts are challenging. Thus, manipulating text cohesion should be taken advantage of and used towards this ultimate goal.

6.4.3 Selecting or Developing Suitable Materials in Terms of Cohesion for EFL Readers

From the multiple studies reviewed in this dissertation, it is revealed that the cohesion of the text alters EFL readers' comprehension and processing. Based on these findings, it is important for teachers to ensure that the English text is of appropriate difficulty and cohesion for the learners. There are many tools that can automatically calculate the characteristics of an English word or sentence by simply copying and pasting them (e.g., Coh-Metrix, a tool for the automatic analysis of cohesion: TACOO, CEFR-based Vocabulary Level Analyzer: CVLA, developed by McNamara et al., 2014, Crossley et al., 2016, and Uchida & Negishi, 2018, respectively). Although analyzing all the English texts taught in a classroom might not be practical, such tools can be used to check the difficulty level of a text that the students struggle with or the material that a teacher finds.

It has been proposed that when texts are technical and complicated, readers may face difficulty reading; thus, increasing text cohesion is helpful (Ozuru et al., 2009). However, similar to the college-level participants of this study, if readers have the ability to understand the textual contents, the texts are not required to be highly cohesive. Rather, low-cohesion texts are more facilitative to make inferences about the text. This is because high-knowledge readers can fill in the gaps in the text when the text content is familiar to

them (McNamara, 2001; McNamara et al., 1996; McNamara et al., 2011).

Specifically, learners in the EFL environment often have problems understanding the message and coherence of a text as a whole (e.g., Morishima, 2013; Ushiro et al., 2020), even when they can translate the individual meanings of words and phrases. A way to help such learners is to confirm that the English texts they are provided with are designed to be reasonably comprehensible to learners.

If the teacher feels that the level does not match the learner, they can modify the text to make it more suitable by, for example, changing pronouns by replacing them with nouns, or by underlining words that refer to the same thing to enhance the input. Alternatively, teachers can provide supplementary instruction in the classroom to help readers connect sentences or incorporate inferential questions that allow readers to infer what is not explicitly stated in the text.

Furthermore, it is necessary to consider the developmental stage of the learner and the nature of the topic or genre of the material being used. In language learning, it is good to give learners something slightly challenging to engage with. Teachers should consider what to focus on in their reading instruction and be careful not to use materials that are too demanding for the readers.

Teachers should care about text cohesion when their students are unable to read their target texts. Although this study focused on college-level students, past studies have shown a positive effect through increasing the text cohesion for relatively novice learners (Beck et al., 1991; Loxterman et al., 1994). Hosoda (2016), who conducted an experimental study on Japanese EFL learners, also discussed the need for high-cohesion text for less skilled learners who have difficulty interpreting meanings and making inferences. Thus, it is preferable that teachers of junior high and senior high schools expressly shape and support students by adding supplementary information (i.e.,

connectives, causal information) to create high-cohesion text and help students grasp the relationships at play in the texts. However, considering that the reverse cohesion effect was also observed in this study, teachers may not need to be too sensitive about the cohesion level of the text they use to teach students English when readers have sufficient proficiency to read the text or the background knowledge to understand the text's contents. Hence, it is suggested that not only teachers but also the creators of teaching materials need to consider the text cohesion along with the traditional difficulty indicators (vocabulary, length, and the traditional difficulty indicator, Flesch-Kincaid Grade Level). It is necessary to check whether the degree of reference cohesion is appropriate for the level of the reader.

6.5 Concluding Remarks

This study shows that text cohesion tends to be weakened in more difficult texts in English language materials for Japanese EFL learners, and that the degree of text cohesion affects readers' processing and comprehension. The present study also revealed that the positive/negative effects of cohesion manipulation on readers' comprehension differed depending on their reading proficiency levels and the text's genre.

Since text cohesion can be one of the difficulty indexes of English texts, EFL teachers need to pay attention to ensure that the presentation of textual information to readers is not overly difficult (low-cohesion texts) and to gauge whether readers find it difficult to make connections between information to form a coherent understanding. Moreover, if readers' proficiency is high enough, providing challenging and authentic texts is recommended.

In addition, material developers need to consider not only vocabulary- and grammar-based graded levels, but also readability considering text cohesion–explicitness,

redundancy, and ease of interpreting elements in the texts. As text cohesion reflects discourse-level difficulty in reading, it would be a useful index to consider EFL readers' processing burden.

I hope this study bridges theory and practice in EFL reading. To better understand EFL reading mechanisms, features, and the effects of text cohesion on EFL reading, more detailed, future research should expand this study and investigate readers' individual factors, text genres, and the relation of these factors to text cohesion in English.

References

Aidinlou, N. A., Khodamard, N., & Azami, J. (2012). The effect of textual cohesive reference instruction on the reading comprehension of Iranian EFL students. *International Journal of English Linguistics*, *2*(5), 18–24. https://doi.org/10.5539/ijel.v2n5p18

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*(5), 1061–1070. https://doi.org/10.1037/0278-7393.19.5.1061

Beck, I. L., McKeown, M. G., Sinatra, G. M., & Loxterman, J. A. (1991). Revising social studies text from a text-processing perspective: Evidence of improved comprehensibility. *Reading Research Quarterly*, *26*(3), 251–276. https://www.jstor.org/stable/747763

Best, R., Ozuru, Y., Floyd, R. G., & McNamara, D. S. (2006). Children's text comprehension: Effects of genre, knowledge, and text cohesion. *ICLS '06 Proceedings of the 7th International Conference on Learning Sciences*, 37–42. http://dl.acm.org/citation.cfm?id=1150040

Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing*, *11*, 489–503. http://dx.doi.org/10.1023/A:1008084120205

Carrell, P. L. (1982). Cohesion is not coherence. *TESOL Quarterly*, *16*(4), 479–488. https://doi.org/10.2307/3586466

Carrell, P. L. (1987). Content and formal schemata in ESL reading. *TESOL Quarterly*, *21*(3), 461–481. https://doi.org/10.2307/3586498

Carrell, P. L., & Eisterhold, J. C. (1983). Schema theory and ESL reading pedagogy. *TESOL Quarterly*, *17*(4), 553–573. https://doi.org/10.2307/3586613

Chen, X., Dronjic, V., & Helms-Park, R. (2015). *Reading in a second language: Cognitive and psycholinguistic issues*. Routledge.

Crossley, S. A., Allen, D., & McNamara, D. S. (2012). Text simplification and comprehensible input: A case for an intuitive approach. *Language Teaching Research*, *16*, 89–108. https://doi.org/10.1177/1362168811423456

Crossley, S., Clevinger, A., & Kim, Y. (2014). The role of lexical properties and cohesive devices in text integration and their effect on human ratings of speaking proficiency. *Language Assessment Quarterly*, *11*(3), 250–270. https://doi.org/10.1080/15434303.2014.926905

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods* *48*(4), 1227–1237. https://doi.org/10.3758/s13428-015-0651-7

Crossley, S. A., Louwerse, M. M., McCarthy, P. M., & McNamara, D. S. (2007). A linguistic analysis of simplified and authentic texts. *The Modern Language Journal*, *91*, 15–30. https://doi.org/10.1111/j.1540-4781.2007.00507.x

Dahl, A., Carlson, S., Renken, M., & Reynolds, E. (2020). *Exploring deep and referential cohesion and its effects on adolescent readers' comprehension processing*. EasyChair. https://easychair.org/publications/preprint/Ncqm

Davis, J. N., & Bistodeau, L. (1993). How do L1 and L2 reading differ? Evidence from thinking aloud protocols. *Modern Language Journal, 77*(4), 459–472. https://doi.org/10.2307/329671

Désiron, J. C., Bétrancourt, M., & de Vries, E. (2020). Cross-representational signaling and cohesion support inferential comprehension of text-picture documents. *Frontiers in Psychology*, *11*, 1–11. https://doi.org/10.3389/fpsyg.2020.592509

Farrokh, P., & Gavabari, S. H. (2019). The study of the effectiveness of textual cohesion of teaching materials on Iranian intermediate EFL learners' reading comprehension. *Journal of Literature, Languages and Linguistics*, *52*, 15–22. https://doi.org/10.7176/JLLL/52-03

Freebody, P., & Anderson, R. C. (1983). Effects of vocabulary difficulty, text cohesion, and schema availability on reading comprehension. *Reading Research Quarterly*, *18*(3), 277–294. https://doi.org/10.2307/747389

Gakken Education Publishing. (2014) *Eiken ni kyu kako mondai-syu*. [The collection of previous questions for the second frade STEP test]. Author.

Gerrig, R. J., & McKoon, G. (1998). The readiness is all: The functionality of memory-based text processing. *Discourse Processes*, *26*(2–3), 67–86. https://doi.org/10.1080/01638539809545039

Gernsbacher, M. A. (1990). *Language comprehension as structure building.* Lawrence Erlbaum Associates.

Gilabert, R., Martínez, G., & Vidal-Abarca, E. (2005). Some good texts are always better: Text revision to foster inferences of readers with high and low prior background knowledge. *Learning and Instruction*, *15*(1), 45–68. https://doi.org/10.1016/j.learninstruc.2004.12.003

Goldman, S. R., & Saul, E. U. (1990). Flexibility in text processing: A strategy competition model. *Learning and Individual Differences*, *2*(2), 181–219. https://doi.org/10.1016/1041-6080(90)90022-9

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

Grabe, W., & Stoller, F. L. (2019). *Teaching and researching reading* (3rd ed.). Routledge.

Graesser, A. C., & McNamara, D. S. (2011). Computational analyses of multilevel discourse comprehension. *Topics in Cognitive Science*, *3*(2), 371–398. https://doi.org/10.1111/j.1756-8765.2010.01081.x

Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher*, *40*(5), 223–234. https://doi.org/10.3102/0013189X11413260

Graesser, A. C., McNamara, D. S., & Louwerse, M. M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text? In A. P. Sweet & C. E. Snow (Eds.), *Rethinking reading comprehension* (pp. 82–98). Guilford.

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*(2), 193–202. https://doi.org/10.3758/BF03195564

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395. https://doi.org/10.1037/0033-295X.101.3.371

Hall, S. S., Basran, J., Paterson, K. B., Kowalski, R., Filik, R., & Maltby, J. (2014). Individual differences in the effectiveness of text cohesion for science text comprehension. *Learning and Individual Differences*, *29*, 74–80. https://doi.org/10.1016/j.lindif.2013.10.014

Hall, S. S., Kowalski, R., Paterson, K. B., Basran, J., Filik, R., & Maltby, J. (2015). Local text cohesion, reading ability and individual science aspirations: Key factors influencing comprehension in science classes. *British Educational Research Journal*, *41*(1), 122–142. https://doi.org/10.1002/berj.3134

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English.* Routledge.

Han, S., & Shin, J. (2016). Exploring the role of automatically-derived text complexity features in L2 reading test development. *The Modern English Education Society*, *17*(2), 1–19.

Helder, A., Van Leijenhorst, L., & van den Broek, P. (2016). Coherence monitoring by good and poor comprehenders in elementary school: Comparing offline and online measures. *Learning and Individual Differences*, *48*, 17–23. https://doi.org/10.1016/j.lindif.2016.02.008

Horiba, Y. (1996). The role of elaborations in L2 text memory: The effect of encoding task on recall of causally related sentences. *The Modern Language Journal*, *80*(2), 151–164. https://doi.org/10.1111/j.1540-4781.1996.tb01155.x

Horiba, Y. (2000). Reader control in reading: Effects of language competence, text type, and task. *Discourse Processes*, *29*(3), 223–267. https://doi.org/10.1207/S15326950dp2903_3

Horiba, Y. (2013). Task-induced strategic processing in L2 text comprehension. *Reading in a Foreign Language*, *25*(2), 98–125. https://eric.ed.gov/?id=EJ1015753

Horiba, Y. (2017). Effects of linguistic knowledge, cohesion, and L1 background on comprehension of L2 expository text. *Studies in Language Sciences*, *23*, 1–20. https://kuis.repo.nii.ac.jp/?action=pages_view_main&active_action=repository_view_main_item_detail&item_id=1450&item_no=1&page_id=13&block_id=17

Hosoda, M. (2016). The interplay of text cohesion and L2 reading proficiency in different levels of text comprehension among EFL readers. *ARELE: Annual Review of English Language Education in Japan*, *27*, 201–216. https://doi.org/10.20581/arele.27.0_201

Hyönä, J., & Nurminen, A. M. (2006). Do adult readers know how they read? Evidence from eye movement patterns and verbal reports. *British Journal of Psychology, 97*(1), 31–50. https://doi.org/10.1348/000712605X53678

Ikeno, O. (1996). The effects of text-structure-guiding questions on comprehension of texts with varying linguistic difficulties. *JACET Bulletin*, *27*, 51–68. http://ci.nii.ac.jp/naid/110003814657

Japan Association of College English Teachers (JACET) Basic Word Revision Committee. (2003). *JACET list of 8000 basic words*. JACET.

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Allyn & Bacon.

Kamalski, J., Sanders, T., & Lentz, L. (2008). Coherence marking, prior knowledge, and comprehension of informative and persuasive texts: Sorting things out. *Discourse Processes*, *45*(4–5), 323–345. https://doi.org/10.1080/01638530802145486

Kasashima, J. (Ed.). (2019a). *New horizon: English course 1*. Tokyoshoseki.

Kasashima, J. (Ed.). (2019b). *New horizon: English course 2*. Tokyoshoseki.

Kasashima, J. (Ed.). (2019c). *New horizon: English course 3*. Tokyoshoseki.

Kimura, Y. (2015). Task instructions for global coherence in EFL reading: Evidence from think-aloud protocols and a recall task. *ARELE: Annual Review of English Language Education in Japan*, *26*, 173–188. https://doi.org/10.20581/arele.26.0_173

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. https://doi.org/10.1037/0033-295X.95.2.163

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*(4), 294–303. https://doi.org/10.1037/0003-066X.49.4.294

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

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363–394. https://doi.org/10.1037/0033-295X.85.5.363

Kitajima, R. (1997). Referential strategy training for second language reading comprehension of Japanese texts. *Foreign Language Annals*, *30*(1), 84–97. https://doi.org/10.1111/j.1944-9720.1997.tb01319.x

Kiyota, Y. (Ed.). (2019a). *All aboard! English communication I*. Tokyoshoseki.

Kiyota, Y. (Ed.). (2019b). *All aboard! English communication II*. Tokyoshoseki.

Kiyota, Y. (Ed.). (2019c). *All aboard! English communication III*. Tokyoshoseki.

Kleijn, S., Pander Maat, H. L., & Sanders, T. J. (2019). Comprehension effects of connectives across texts, readers, and coherence relations. *Discourse Processes, 56*(5–6)*,* 447－464. https://doi.org/10.1080/0163853X.2019.1605257

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

Leow, R. P., & Morgan-Short, K. (2004). To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition*, *26*(1), 35–57. https://doi.org/10.1017/S0272263104026129

Li, M., & D'Angelo, N. (2015). Higher-level process in second language reading comprehension . In X. Chen, V. Dronjic, & R. H. Park (Eds.), *Reading in a second language: Cognitive and psycholinguistic issues* (pp. 158–194). Routledge.

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*(4), 525–556. https://doi.org/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 texts. *Journal of Educational Psychology*, *94*(4), 778–784. https://doi.org/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*(2), 165–186. https://doi.org/10.1207/s15326950dp3702_5

Loxterman, J. A., Beck, I. L., & McKeown, M. G. (1994). The effects of thinking aloud during reading on students' comprehension of more or less coherent text. *Reading Research Quarterly*, *29*(4), 353–367. https://doi.org/10.2307/747784

McCrudden, M. T., & McNamara, D. S. (2017). *Cognition in education*. Routledge.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, *99*(3), 440–466. https://doi.org/10.1037/0033-295X.99.3.440

McNamara, D. S. (2001). Reading both high-coherence and low-coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology*, *55*(1), 51–62. https://doi.org/10.1037/h0087352

McNamara, D. S., & Kintsch, W. (1996). Learning from text: Effects of prior knowledge and text coherence. *Discourse Processes*, *22*(3), 247–287. https://doi.org/10.1080/01638539609544975

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), 1–43. https://doi.org/10.1207/s1532690xci1401_1

McNamara, D. S., Graesser, A. C., & Louwerse, M. M.(2012). Sources of text difficulty: Across genres and grades. In J. P. Sabatini, E. Albro, & T. O'Reilly (Eds.), *Measuring up: Advances in how we assess reading ability* (pp. 89–116). Rowman & Littlefield Education.

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 297–384). Elsevier Academic Press. https://doi.org/10.1016/S0079-7421(09)51009-2

McNamara, D. S., Ozuru, Y., & Floyd, R. G. (2011). Comprehension challenges in the fourth grade: The roles of text cohesion, text genre, and readers' prior knowledge. *International Electronic Journal of Elementary Education*, *4*(1), 229–257. https://www.iejee.com/index.php/IEJEE/article/view/222

Ministry of Education, Culture, Sports, Science, and Technology (Ed.). (2018a). *We can! 1*. Tokyoshoseki.

Ministry of Education, Culture, Sports, Science, and Technology (Ed.). (2018b). *We can! 2*. Tokyoshoseki.

Mizumoto, A. (2013). Eibunnkaisekipuroguramu kara erareru kakusyusihyou wo tsukatta tekisutonannido no suitei: Kyouzaisakusei heno tekiyoukanousei [Estimation of text difficulty using various indices from English analysis programs: Applicability to the preparation of teaching materials]. *LET Kansai Chapter Collected Pape*r, 141–150.

Mokhtari, K., & Sheorey, R. (2002). Measuring ESL students' awareness of reading strategies. *Journal of Developmental Education*, *25*(3), 2–10. Retrieved from https://eric.ed.gov/?id=EJ645740

Mori, Y. (2018). Syougakkoeigokyouiku ni okeru sinnkyuukyouzai no sutoorii no bunnseki–Tango, bunn, dannwa reberu no yomiyasusa no kanntenn kara [An analysis of stories in new and old materials for teaching English in elementary schools: From the perspective of readability at the word, sentence, and discourse levels]. *Tsukuba Review of English Language Teaching*, *39*, 121–133.

Morishima, Y. (2013). Allocation of limited cognitive resources during text comprehension in a second language. *Discourse Processes*, *50*(8), 577–597. https://doi.org/10.1080/0163853X.2013.846964

Myers, J., & O'Brien, E. (1998). Accessing the discourse representation during reading. *Discourse Processes*, *26*(2–3), 131–157. https://doi.org/10.1080/01638539809545042

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language*, *26*(4), 453–465. https://doi.org/10.1016/0749-596X(87)90101-X

Nahatame, S. (2013). Predictive inference generation during Japanese EFL reading: Focusing on contextual constraint and local coherence of text. *JACET (The Japan Association of College English Teachers) Journal, 56*, 19–38.

Nahatame, S. (2017). Standards of coherence in second language reading: Sentence connectivity and reading proficiency. *Reading in a Foreign Language*, *29*, 86–112. https://files.eric.ed.gov/fulltext/EJ1137893.pdf

Nahatame, S. (2018). Shougakuseimuke sutooriikyouzai no bunnmyaku no bunnseki [Analyzing the contexts of English story materials for Japanese elementary school students: An investigation based on semantic relatedness between sentences]. *JES Journal*, *18*, 84–99. https://doi.org/10.20597/jesjournal.18.01_84

Nahatame, S. (2021). Nennsyoumuke eigoyomimonokyouzai niokeru bunnsyo no kessokusei [Cohesion features of material for young readers of English: A computational investigation]. *The Science of Reading*, *62*(3–4), 146–159. https://doi.org/10.19011/sor.62.3-4_146

Nahatame, S., & Kimura, Y. (2019). Gengotekitokutyou ni motozuku syougakuseimuke sutooriikyouzai no bunnrui: Eibunnkaisekipuroguramu wo mochiita tamenntekina bunnseki [Clustering English story materials for Japanese elementary school students based on their linguistic features: A multidimensional approach using a text mining tool]. *JES Journal*, *19*, 70–85. https://doi.org/10.20597/jesjournal.19.01_70

Namai, K. (Ed.). (2019a). *New discovery English communication 1*. Kairyudo.

Namai, K. (Ed.). (2019b). *New discovery English communication 2*. Kairyudo.

Namai, K. (Ed.). (2019c). *New discovery English communication 3*. Kairyudo.

Negishi, M. (Ed.). (2016a). *New crown 1*. Sanseido.

Negishi, M. (Ed.). (2016b). *New crown 2*. Sanseido.

Negishi, M. (Ed.). (2016c). *New crown 3*. Sanseido.

Niisato, M. (Ed.). (2016a). *Sunshine: English course 1*. Kairyudo.

Niisato, M. (Ed.). (2016b). *Sunshine: English course 2*. Kairyudo.

Niisato, M. (Ed.). (2016c). *Sunshine: English course 3*. Kairyudo.

Oakhill, J. (1984). Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology*, *54*(1), 31–39. https://doi.org/10.1111/j.2044-8279.1984.tb00842.x

Obunsha. (2003a). *Eiken jyun ni kyu zen mondai-syu*. [The collection of previous questions for the pre-second grade EIKEN test]. Obunsha.

Obunsha. (2003b). *Eiken ni kyu zen mondai-syu*. [The collection of previous questions for the pre-second grade EIKEN test]. Obunsha.

Obunsha. (2006). *Eiken jyun ni kyu zen mondai-syu*. [The collection of previous questions for the pre-second grade EIKEN test]. Obunsha.

Obunsha. (2009). *Eiken jyun ichi kyu zen mondai-syu*. [The collection of previous questions for the pre-first grade EIKEN test]. Obunsha.

Obunsha. (2014). *Eiken jyun ichi kyu zen mondai-syu*. [The collection of previous questions for the pre-first grade EIKEN test]. Obunsha.

Oh, S. Y. (2001). Two types of input modification and EFL reading comprehension: Simplification versus elaboration. *TESOL Quarterly*, *35*(1), 69–96. https://doi.org/10.2307/3587860

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*(2), 121–152. https://doi.org/10.1080/01638530709336895

Ozono, S., & Ito, H. (2003). Logical connectives as catalysts for interactive L2 reading. *System*, *31*, 283–297.

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*(8), 641–667. https://doi.org/10.1080/01638531003628809

Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction*, *19*(3), 228–242. https://doi.org/10.1016/j.learninstruc.2008.04.003

Plakans, L., & Zeynep, B. (2016). Cohesion features in ESL reading: Comparing beginning, intermediate and advanced textbooks. *Reading in a Foreign Language*, *28*(1), 79–100. https://scholarspace.manoa.hawaii.edu/handle/10125/66716

Radvansky, G. A., Tamplin, A. K., Armendarez, J., & Thompson, A. N. (2014). Different kinds of causality in event cognition. *Discourse Processes*, *51*(7), 601–618. https://doi.org/10.1080/0163853X.2014.903366

Rapp, D. N., & Kendeou, P. (2007). Revising what readers know: Updating text representations during narrative comprehension. *Memory & Cognition*, *35*, 2019–2032. dhttps://doi.org/10.3758/BF03192934

Roman, D. X., Briceño, A., Rohde, H., & Hironaka, S. (2016). Linguistic cohesion in middle-school texts: A comparison of logical connectives usage in science and social studies textbooks. *The Electronic Journal for Research in Science & Mathematics Education*, *20*(6), 1–19. https://ejrsme.icrsme.com/article/view/15997

Rouet, J.-F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19–52). IAP Information Age Publishing.

Safaie, S. (2020). The effects of explicit and implicit teaching of connectors on the reading comprehension performance of Iranian EFL learners. *Cogent Education*, *7*(1), 1–18. https://doi.org/10.1080/2331186X.2020.1777806

Sanders, T. J., Spooren, W. P., & Noordman, L. G. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, *15*(1), 1–35. https://doi.org/10.1080/01638539209544800

Schmitz, A., Gräsel, C., & Rothstein, B. (2017). Students' genre expectations and the effects of text cohesion on reading comprehension. *Reading and Writing, 30*(5), 1115–1135. https://doi.org/10.1007/s11145-016-9714-0

Shimizu, H. (2009). The effects of causal relatedness on EFL learners' reading comprehension and inference generation. *ARELE: Annual Review of English Language Education in Japan*, *20*, 31–40. https://doi.org/10.20581/arele.20.0_31

Shimizu, H. (2015). Generation of local and global inferences in L2 reading comprehension. *JACET Journal*, *59*, 75–92. https://ci.nii.ac.jp/naid/110009907979

Smith, R., Snow, P., Serry, T., & Hammond, L. (2021). The role of background knowledge in reading comprehension: A critical review. *Reading Psychology*, *42*(3), 214–240. https://doi.org/10.1080/02702711.2021.1888348

Tabata, Y., & Masuda, M. (2018). Nihon no syougakkogaikokugokatudouhukudokuhonn to finlando no kisogakkoeigokyoukasyo no hikaku to kennto– Eigogakusyuukaisidannkai ni tyakumokushite– [A comparative study on English textbook in Japan and Finland: Focus on English learning start stage]. *Departmental Bulletin Paper in Faculty of Human Development in Toyama University*, *13*, 81–93.

Takaki, S. (2014). *Modeling of EFL reading comprehension based on the event-indexing model focusing on the role of situational continuity* [Unpublished doctoral dissertation]. University of Tsukuba, Japan.

Tanabe, M. (Ed.). (2019a). *Prominence: English communication I*. Tokyoshoseki.

Tanabe, M. (Ed.). (2019b). *Prominence: English communication II*. Tokyoshoseki.

Tanabe, M. (Ed.). (2019c). *Prominence: English communication III*. Tokyoshoseki.

Todaro, S., Millis, K., & Dandotkar, S. (2010). The impact of semantic and causal relatedness and reading skill on standards of coherence. *Discourse Processes*, *47*(5), 421–446. https://doi.org/10.1080/01638530903253825

Torikai, S., Kumagai, M., Hirata, M., Yoshida, Y., & Kobayashi, M. (2017). Doitu to nihon no tyuutoukyouiku de siyousareteiru eigokyoukasyo no hikakubunnseki [A comparative analysis of English textbooks used in German and Japanese secondary education]. *Intercultural Communication Review*, *15*, 63–78.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, *24*(5), 612–630. https://doi.org/10.1016/0749-596X(85)90049-X

Traxler, M. J. (2011). *Introduction to psycholinguistics: Understanding language science*. John Wiley & Blackwell.

Uchida, S. & Negishi, M. (2018) Assigning CEFR-J levels to English texts based on textual features. In Y. Tono & H. Isahara (Eds.), *Proceedings of the 4th Asia Pacific Corpus Linguistics Conference (APCLC 2018)* (pp. 463–467).

Ushiro, Y., Hijikata, Y., Shimizu, M., Nakagawa, C., Koga, T., Ohno, M., & Umehara, C. (2007). Relationships between cue types in recall tests and L2 reading proficiency. *ARELE: Annual Review of English Language Education in Japan*, *18*, 31–40. https://doi.org/10.20581/arele.18.0_31

Ushiro, Y., Hamada, A., Mori, Y., Hosoda, M., Tada, G., Kamimura, K., & Okawara, N. (2018). Goal-oriented L2 reading processes in maintaining the coherence of narrative comprehension. *JACET Journal*, 62, 109–128. https://doi.org/10.32234/jacetjournal.62.0_109

Ushiro, Y., Ogiso, T., Nahatame, S., Hosoda, M., Kamimura, K., Sasaki, Y., Komuro, R., Okada, R., & Aoki, S. (2020). EFL readers' understanding of protagonist, temporal, and spatial links in narrative: Evidence from eye tracking. *ARELE: Annual Review of English Language Education in Japan*, *31*, 97–112. https://doi.org/10.20581/arele.31.0_97

van den Broek, P. (1994). Comprehension and memory of narrative text. In M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 539–588). Academic Press.

van den Broek, P., Lorch, R. F. Jr., Linderholm, T., & Gustafson, M., (2001). The effect of readers' goals on the generation of inferences. *Memory & Cognition*, *29*(8), 1081–1087. https://doi.org/10.3758/BF03206376

van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Processes*, *39*(2–3), 299–316. https://doi.org/10.1080/0163853X.2005.9651685

van den Broek, P., Risden, K., Fletcher, C. R., & Thurlow, R. (1996). A "landscape" view of reading: Fluctuating patterns of activation and the construction of a stable memory representation. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 165–187). Lawrence Erlbaum Associates, Inc.

van den Broek, P., Risden, K., & Husebye-Hartmann, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In R. F. Lorch, Jr., & E. J. O'Brien (Eds.), *Sources of coherence in text comprehension* (pp. 353–373). Lawrence Erlbaum Associates, Inc.

van den Broek, P., Virtue, S., Everson, M. G., Tzeng, Y., & Sung, Y. (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). Lawrence Erlbaum Associates.

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and the online construction of a memory representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71–98). Erlbaum.

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

van Silfhout, G., Evers-Vermeul, J., & Sanders, T. (2015). Connectives as processing signals: How students benefit in processing narrative and expository texts. *Discourse Processes*, *52*(1), 47–76. https://doi.org/10.1080/0163853X.2014.905237

Vidal-Abarca, E., Martínez, G., & Gilabert, R. (2000). Two procedures to improve instructional text: Effects on memory and learning. *Journal of Educational Psychology*, *92*(1), 107–116. 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*(1), 45–68. https://doi.org/10.1207/s1532690xci1401_2

Warren, P. (2012). *Introducing psycholinguistics*. Cambridge University Press.

Yano, Y., Long, M. H., & Ross, S. (1994). The effects of simplified and elaborated texts on foreign language reading comprehension. *Language Learning*, *44*(2), 189–219. https://doi.org/10.1111/j.1467-1770.1994.tb01100.x

Yuill, N., & Oakhill, J. (1991). *Children's problems in text comprehension: An experimental investigation*. Cambridge University Press.

Zhang, J. (2018). The effect of strategic planning training on cohesion in EFL learners' essays. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, *34*(6), 554–567. https://doi.org/10.1080/10573569.2018.1506371

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, *6*(5), 292–297. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x

Appendices

Appendix A: The High-Cohesion Heat/EH Text Used in Experiments 1, 3, and 4

Effects of Heat on Objects and Matter

Heat Moves

Heat can move from one object to another object, or it can move from one place to another place. Heat moves from warm objects to cooler ones. For example, you can warm your cold hands by holding a cup of warm soup. Your hands become warmer because heat moves from the soup, through the cup, to your hands. The heat from the soup also moves above the cup. Thus, you can feel warm air rising above the cup.

Heat moves through some materials more easily than other materials. Conductors[1] are materials through which heat moves easily. Most metals are good conductors. For example, metal pots are used for cooking because heat from the stove[2] quickly moves through the metal pots and the heat in the pot warms the food.

Heat Changes Matter

Adding heat or taking away heat can change matter. Matter is something that takes up space. Matter can change from one state to another state, or from one form to another form. Three states of matter are solid, liquid and gas. For example, an ice cube is solid water. Heat can melt an ice cube, causing the ice cube to change into liquid water. When heat is taken away, the liquid water can change back into solid water (ice).

Heat can make liquids boil and change into a gas state. For example, water boils when it is heated. As the water boils, it turns into a gas state that is called water vapor[3]. Heat from the sun causes liquid water to turn into water vapor. Water vapor then mixes with the air in a process called evaporation[4].

However, sometimes heat causes changes that cannot be changed back. As one example, bread can change into toast when you heat the bread. However, you cannot untoast a piece of toast by taking away heat. As another example, eggs change when you cook them in a pan. However, of course you cannot uncook an egg by taking away the heat.

---

[1] 伝導体

[2] コンロ

[3] 蒸気

[4] 蒸発

Effects of Heat

Moving Heat

Heat can move from one object or place to another. Heat moves from warm objects to cooler ones. You can warm your hands by holding a cup of warm soup. Heat moves from the soup through the cup to your hands. You can feel warm air rising above the cup.

Heat moves through some materials more easily than others. Heat moves easily through conductors[5]. Most metals are good conductors. Metal pots are used for cooking. Heat from the stove[6] quickly moves through the metal. The heat warms the food.

Changing Matter

Adding or taking away heat can change matter. Matter is something that takes up space. Matter can change from one state, or form, to another. An ice cube is solid water. Solid is one state of matter. Heat can melt an ice cube. The ice cube changes into liquid water. Liquid is another state of matter. When heat is taken away, the water can change back. Liquid water turns into solid water.

Heat can make liquids boil. Water boils when it is heated. When the water boils, it turns into a gas. This gas is called water vapor[7]. Solid, liquid and gas are three states of matter. Heat from the sun causes liquid water to turn into water vapor. Water vapor mixes with the air. This is called evaporation[8].

Sometimes heat causes changes that cannot be changed back. Bread can change into toast when you heat it. Eggs change when you cook them in a pan. You cannot untoast a piece of toast. You cannot uncook an egg.

[5] 伝導体

[6] コンロ

[7] 蒸気

[8] 蒸発

Appendix C: The High-Cohesion Space Text Used in Experiment 1

Space Travel and the Human Body

While in space, the body is not affected by gravity. Therefore, blood and water do not travel to the lower part 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.

Appendix D: The Low-Cohesion Space Text Used in Experiment 1

Space Travel and the Human Body

While in space, the body is not affected by gravity. Blood and water do not travel to the lower part of the body, especially the legs. Instead, 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, their 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.

Appendix E: The Cohesion-Judgment Task Used in Experiment 1

【結束性判断課題】

先ほど読んだ英文は，内容がはっきりと，わかりやすく関連付けて記述されていた

1: 全く思わない

2: 思わない

3: どちらかといえばそう思わない

4: どちらかと言えばそう思う

5: そう思う

6: とてもそう思う

Appendix F: The Reading Strategy Questionnaire Used in Experiment 1

みなさんが英語教材を読む (例 宿題や試験のために教科書を読む) 時にどのようなことに気をつけて英語を読んでいるかを調査するアンケートです。各質問に関して、最も当てはまる数字を選び、当てはまる数字に丸をつけて下さい。

質問1. テキストを読む時、目的を持つようにしている (例 単語を学ぶため、調べ学習をするため)。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問2. テキストの内容を理解するために、読む時にメモを取っている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問3. テキストの内容を理解するために、テキストに関連して自分が知っていることを考えるようにしている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問4. 読み始める前に、テキストにどんなことが書かれているかを知るため、全体に目を通している。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問5. テキストが難しいと感じた時には、声に出して文章を読み、読んでいる部分を理解しようとしている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問6. テキスト内容が自分の読みの目的 (例 単語を学ぶ、調べ学習) に合っているかどうかを考えている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問7. テキストの内容を理解するために、時間をかけて注意深く読んでいる。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問8.　はじめにテキストを読む時には、テキストの長さや全体の構成などの特徴に注意している。

1	2	3	4	5
全く行わない	ごくたまに行う	時々（50%）行う	だいたいいつも行う	必ずいつも行う

質問9.　集中力がなくなった時には、テキストの本題に戻るようにしている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々（50%）行う	だいたいいつも行う	必ずいつも行う

質問10.　テキストの内容を覚えるために、テキストに下線を引いたり、丸で囲んだりしている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々（50%）行う	だいたいいつも行う	必ずいつも行う

質問11.　読んでいる内容によって、読むスピードを変えている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々（50%）行う	だいたいいつも行う	必ずいつも行う

質問12.　テキストを読んでいる時に、どの情報を注意深く読むか・どの情報を無視するかを自分で決めている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々（50%）行う	だいたいいつも行う	必ずいつも行う

質問13.　テキスト内容を理解するために、辞書などを使っている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々（50%）行う	だいたいいつも行う	必ずいつも行う

質問14.　テキストが難しいと感じた時には、読んでいる内容により注意を払うようにしている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々（50%）行う	だいたいいつも行う	必ずいつも行う

質問15.　テキストをよく理解するために、テキスト内の表や図、写真などを活用している。

1	2	3	4	5
全く行わない	ごくたまに行う	時々（50%）行う	だいたいいつも行う	必ずいつも行う

質問16.　読んでいる内容について考えるために、読むのを一度やめて、内容について考えている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々（50%）行う	だいたいいつも行う	必ずいつも行う

質問17. テキストの内容をよりよく理解するために、読んでいる時には周りの手掛かりを利用している。(例 分からない単語があったときに周りの内容から考える)

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問18. テキストの内容をより良く理解するために、自分の言葉を使ってテキストの内容を言い換えている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問19. テキスト内容を覚えるために、テキスト内容の情景やイメージを頭に思い浮かべている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問20. テキストの中で大事な情報を見つけるために、太字や斜体などの、テキストの印刷上の特徴に注意している。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問21. テキストで言われている情報が正しいか、批判的に考えたり、評価したりしている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問22. テキスト中の文どうしの関係 (例 主張と具体例) を見つけるために、読み返したり、先読みしたりしている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問23. テキストを読んでいる時に新しい情報に出会った時には、自分がそれまでの内容を理解できているか確かめている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問24. テキストを読む時には、先の内容がどんなものであるか予測しながら読んでいる。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問26. テキストが難しいと感じた時には、理解を深めるために、一度読んだ部分をもう一度読むようにしている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問27. 読むテキストでどんなことを知りたいか、自分自身に問いかけながら読んでいる。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問28. テキストの内容を予測しながら読んでいる時には、自分の予測が正しかったかどうか確かめている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問29. テキストを読んでいる時に出会った知らない単語や熟語の意味を推測して読んでいる。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問30. テキストを読んでいる時、英語を自分の母語 (日本語) に訳しながら読んでいる。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

質問31. テキストを読んでいる時、内容について英語と母語 (日本語) の両方で考えている。

1	2	3	4	5
全く行わない	ごくたまに行う	時々 (50%) 行う	だいたいいつも行う	必ずいつも行う

Fast Food for Bears

We usually think of bears as living in the wild, but these days more and more bears are spending time in towns. A study by the Wildlife Conservation Society of the United States has shown that the main cause of this is the garbage[9] that is found near fast-food restaurants and people's homes. Garbage is a great source of food for bears because bears can always find garbage in the same place all year round, while other food can't always be found in the same place.

The researchers followed 59 bears for 24 hours in the Lake Tahoe[10] area between Nevada and California[11]. The researchers found that there were two types of bears: "country bears," who spent nearly all their time in the wild, and "city bears," who stayed mainly in the towns. Bears are usually active during the day, but the city bears rested during the day and were active at night. While country bears spent a lot of time searching for food, city bears simply ate the never-ending[12] supply of food in people's garbage. The difference between the eating habits of the city and country bears has caused a serious problem. That is, that some city bears have grown up to 270 kilograms ― twice the weight of normal country bears.

Eating human food does not itself seem to harm the bears. Rather, the problem is that towns are very dangerous places for bears. When bears come near humans, bears often get killed. Most bears die in traffic accidents, but an increasing number of bears are killed in other ways. For example, when the bears break into people's homes, wildlife officials often have to shoot bears for safety reasons.

In order to protect both the local communities in the towns where the city bears live and the bears themselves, a campaign called "A Fed Bear Is a Dead Bear" has been started. Campaigners[13] say that the best way to protect the bears is to stop bears from eating garbage by using garbage containers that are too difficult for bears to open. When people participate in this campaign, the bears leave, and this makes life safer for both animals and human beings.

[9] ゴミ
[10] Lake Tahoe　タホ湖
[11] ネバダ州とカリフォルニア州
[12] 尽きることのない
[13] キャンペーンをする人、活動家

Fast Food for Bears

We usually think of bears as living in the wild, but these days more and more of them are spending time in towns. A study by the Wildlife Conservation Society of the United States has shown that the main cause of this is the garbage[14] that is found near fast-food restaurants and people's homes. Garbage is a great source of food for bears because they can find it in the same place all year round.

The researchers followed 59 bears for 24 hours in the Lake Tahoe[15] area between Nevada and California[16]. They found that there were two types of bears: "country bears," who spent nearly all their time in the wild, and "city bears," who stayed mainly in the towns. Bears are usually active during the day, but the city bears rested during the day and were active at night. While country bears spent a lot of time searching for food, city bears simply ate the never-ending[17] supply of food in people's garbage. Some city bears have grown up to 270 kilograms ― twice the weight of normal bears.

Eating human food does not itself seem to harm the bears. The problem is that towns are very dangerous places for them. When bears come near humans, they often get killed. Most die in traffic accidents, but an increasing number of bears are killed in other ways. When the bears break into people's homes, wildlife officials often have to shoot them for safety reasons.

To protect both the local community and the bears, a campaign called "A Fed Bear Is a Dead Bear" has been started. Campaigners[18] say that the best way to protect the bears is to stop them from eating garbage by using garbage containers that are too difficult for them to open. When people do this, the bears leave, and this makes life safer for both animals and human beings.

[14] ゴミ

[15] Lake Tahoe タホ湖

[16] ネバダ州とカリフォルニア州

[17] 尽きることのない

[18] キャンペーンをする人、活動家

Two-for-One

In Peru[19] all children need to attend primary school. However, in reality 23 percent of students stop going to school before the fifth grade. This is often because the children find the classes too difficult. Many of the children come from poor families that cannot even afford to buy books and newspapers. Sometimes their parents themselves cannot read and write. This means that the children have little chance at home to learn the basic skills of reading and writing necessary to understand their classes.

To help these poor children stay in school, UNICEF[20], an organization that works for the right of every child and tries to overcome poverty, is working with the Peruvian[21] government to carry out a new program called Two-for-One. Under this program, first- and second-grade students are taught by teenage[22] volunteers from local high schools. The Two-for-One program targets children from poor families who are doing badly at school. The children meet in groups with the volunteers outside of school hours, and each group selects a study project to work on for eight weeks. These projects help children improve their reading and writing ability while developing social skills through group work. The volunteers try to make this program experience fun so that the children will become more interested in learning and want to stay in school for longer.

The Two-for-One program only started a few years ago, but the program has proved very successful in helping poor children to do better at school. For example, after just eight weeks in the program, children score on average 30 percent higher on school tests. In addition, more than 80 percent of the children who take the program decide to stay in school instead of dropping out. In fact, the program has been so successful that there are now more than 2000 Two-for-One groups active in Peru. The Peruvian government hopes to see this number grow even more in the future.

[19] ペルー(国名)

[20] 国連児童基金、ユニセフ

[21] ペルーの

[22] 十代の

Two-for-One

In Peru[23] all children need to attend primary school. In reality 23 percent of students stop going to school before the fifth grade. They find the classes too difficult. Many of the children come from poor families that cannot even afford to buy books and newspapers. Sometimes their parents themselves cannot read and write. This means that the children have little chance at home to learn the basic skills necessary to understand their classes.

To help these children stay in school, UNICEF[24] is working with the Peruvian[25] government to carry out a new program called Two-for-One. Under this program, first- and second-grade students are taught by teenage[26] volunteers from local high schools. Two-for-One targets children from poor families who are doing badly at school. The children meet in groups with the volunteers outside of school hours, and each group selects a study project to work on for eight weeks. These projects help children improve their reading and writing ability while developing social skills through group work. The volunteers try to make the experience fun so that the children will become more interested in learning.

The Two-for-One program only started a few years ago, but it has proved very successful in helping poor children to do better at school. After just eight weeks in the program, children score on average 30 percent higher on school tests. In addition, more than 80 percent of the children who take the program stay in school. The program has been so successful that there are now more than 2000 Two-for-One groups active in Peru. The Peruvian government hopes to see this number grow even more in the future.

[23] ペルー(国名)

[24] 国連児童基金、ユニセフ

[25] ペルーの

[26] 十代の

Appendix K: The Sheet of the Post-Reading Task Used in Experiment 2

Fast Food for Bears　A　/　B　内容理解タスク　　(氏名)

先ほど読んだ英文に関する問題が 2 問あります。
(ⅰ)→(ⅱ)の順で取り組んでください。その際(ⅱ)を解いた後にもう一度(ⅰ)に戻ることはしないでください。

(ⅰ) 先ほど読んだ英文に関して思い出せることをできるだけ沢山、日本語で書きだしてください。テキストのテーマや詳細情報など、どんな内容でも構いません。ただし、単なる一文ごとの箇条書きや、単語の羅列ではなく、1 つの文章の形で書いてください。

→解答蘭を全て埋める必要はありません。
これ以上書けないと思ったら裏面に進まずに、鉛筆を置き次の指示を待って下さい。

（ⅱ）先ほど読んだ英文に関する以下の質問に**日本語で**答えてください。
・ A Fed Bear Is a Dead Bear キャンペーンが人間と動物の両方にとって良い影響を与えるのはなぜですか。その理由を出来るだけたくさん挙げてください。

→解答欄を全て埋める必要はありません。
これ以上書けないと思ったら鉛筆を置いて次の指示を待って下さい。
また(ⅱ)の回答後に、表面の(ⅰ)に戻ることはできません。

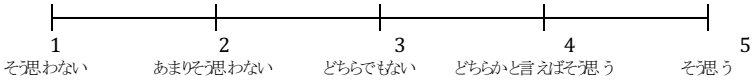
Appendix L: The Questionnaire for FFB Used in Experiment 2

Fast Food for Bears　A　/　B　　アンケート

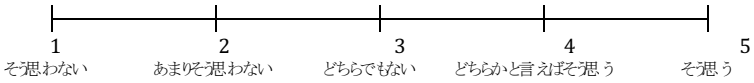
(氏名)＿＿＿＿＿＿＿＿＿＿

これは、先ほど読んだ英文「Fast Food for Bears」に関するアンケートです。
(1)～(8)までの質問に対して、1から5のうち自分に一番当てはまるものに○をしてください。
これはテストではないので解答は自分の主観で構いません。悩まず、直観的に質問に答えて下さい。

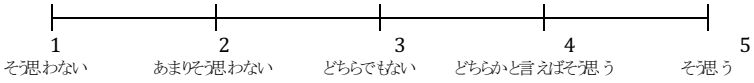
1. この英文を自分はきちんと読めた。



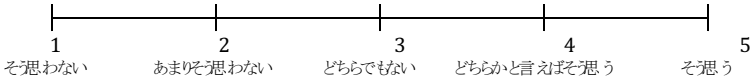
2. この英文は自分にとって読みやすかった。



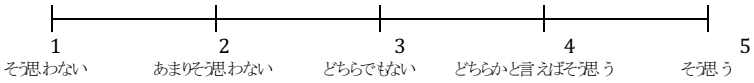
3. この英文を読んでいる際、英文中の出来事や情景が頭にイメージできた。



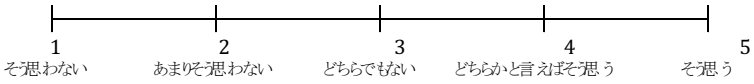
4. この英文には余分な情報はなく必要な情報がはっきりとわかりやすく記述されていた。



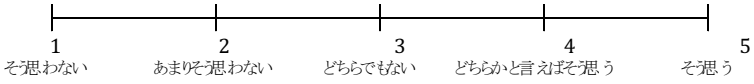
5. この英文中に自分にとって理解しづらい部分があった。



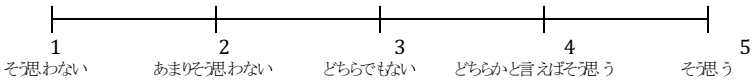
6. この英文を読んでいる際には特に細かい情報にも注意をはらって読んだ。



7. この英文の内容を、読解後に思い出してスムーズに解答できた。



8. この英文で説明されていた内容について、その話の流れはわかりやすかった。



The Needs of Plants

What Plants Need

Plants have certain needs, just like all living things have needs. For example, plants need sunlight, water, and air to live. Plants also need minerals (pronounced[27] as MIN·uhr·uhlz). A mineral is not a plant or an animal. Instead, a mineral is a substance in the ground that occurs naturally. There are three parts of plants that help plants get what they need or help plants make what they need.

The Three Parts of a Plant

The three parts of the plant are the roots, stems, and leaves.

1. The Root

The root is the part of the plant that grows underground. All plants get water and minerals from the ground, which is sometimes called soil. Roots help the plant take in water and minerals that the plant needs from the soil. Roots also help hold the plant in the ground.

2. The Stem

The stem is the part that supports the plant. The stem helps the plant stand upright. It carries minerals and water from the roots of the plant to other parts of the plant. The stem also carries food from the leaves to other parts of the plant.

3. The Leaves

The leaves help the plant make its food. The leaves need sunlight, air, and water to make food. Many leaves have broad, flat surfaces, and these surfaces are broad and flat in order to help the leaves take in lots of sunlight. The energy in sunlight is trapped by the leaf by a substance called *chlorophyll* [28](pronounced KLO ro fill). Leaves are green because chlorophyll is green.

The leaf also helps the plant get the air it needs to make food. This process is helped by tiny holes in each leaf. These holes take in air for the plant. The leaf only uses a gas in the air called *carbon dioxide* (CAR bun di OK side). However, the plant needs both carbon dioxide and water to make food. The plant uses the Sun's energy to combine the carbon dioxide and water to make food. The stem then carries this food to the other parts of the plant.

[27] ～と発音する

[28] クロロフィル、葉緑素

The Needs of Plants

What Are the Needs of Plants?

Like all living things, plants have certain needs. Plants need sunlight, water, and air to live.

Plants also need minerals (MIN·uhr·uhlz). A mineral is a naturally occurring substance that is neither plant nor animal.

The parts of plants help them to get or make what they need. All plants get water and minerals from the soil. The root is the part of the plant that grows underground. Roots help hold the plant in the ground. Roots also help take in water and minerals that the plant needs.

The stem is the part that supports the plant. It helps the plant stand upright. It carries minerals and water from the roots. It also carries food from the leaves to other parts of the plant.

Why Does a Plant Need Leaves?

The leaves (singular[29], *leaf*) are the main food-making part of the plant. Many leaves have broad, flat surfaces that help them take in sunlight. Leaves are green because of *chlorophyll*[30]. Chlorophyll traps the energy (EN·uhr·jee) in sunlight for the plant.

The leaf also helps the plant get the air it needs. Each leaf has tiny holes that take in air for the plant. The leaf uses a gas in the air called *carbon dioxide*. The plant uses the Sun's energy to combine carbon dioxide and water to make food. The stem then carries the food to the other parts of the plant.

[29] 単数形の語

[30] クロロフィル、葉緑素

Appendix O: R Scripts Used in Experiment 3-B

```
K_RECALL<-glmer(RECALL~Coh+(1|Subject)+(1|Item),data=DATA, family = "binomial")
summary(K_RECALL)
```

Note. Coh = Cohesion

Appendix P: The Passage for Practicing Think-Aloud Task Used in Experiment 4

英文を読んでいる際に頭に浮かんできたことを、声に出しながら英文を読んでもらいます。

読解中に行っている単語や文の意味の取り方や、推測したこと、分からないこと、考えていることなどを自由に声に出して読みを進めてください。

◎モデルパッセージ
（モデルの発話を聞きながら、一緒にテキストを追ってみましょう！）

Mammals are also successful because over many, many years, they have developed different kinds of specialized teeth. These different teeth allow mammals to eat many different kinds of food. If mammals can eat many different kinds of food, then they will be less likely to die of starvation[31] and become extinct[32] when a change occurs to one of their food sources.

◎練習用パッセージ
（では、実際にやってみましょう！）
There are four types of teeth in mammals. The number and shape of each of these types of teeth are related to the kind of food the mammal eats. Meat-eating mammals, such as wolves and lions, have long, pointed canine[33] teeth, which are used for cutting. Plant-eating mammals, such as horses and cows, have large, flat premolars and molars[34]. These teeth are used for grinding[35] plant materials. Mammals such as humans have many different kinds of teeth. These help them eat the many different kinds of food in their diets.

[31] 餓死
[32] 絶滅した
[33] 犬歯
[34] 小臼歯と臼歯
[35] すりつぶすこと

Appendix Q: Categories of Think-Aloud Protocols and Examples in Experiment 4

Process level	Category	Definition	Example
Analysis	Word analysis	The reader attempt to analyze the formal or semantic features of a word, phrase, and sentence.	liquid... 液体か
	Sentence analysis		温かい object からそれより冷たい部分に移動する傾向にある
	Reading aloud	The readers' reading aloud that is produced for his/her capturing meaning, following sentence, etc.	Heat moves from warm objects to cooler ones.
Paraphrase	Paraphrase	The readers attempt to paraphrase the expression in the text to enhance his/her understanding.	メタルの台所用品ってことか
Inference	Backward inference	The reader generates an inference that is intended to explain the contents of the current sentence by connecting it to prior text or on the basis of their general knowledge.	あったかいから冷たい空気中にあがるってことかな
	Predictive inference	The reader anticipates something about what will occur in the incoming text.	状態変化とかかな
Response	Association	The reader generates an inference that is brought to in mind to enhance the understanding of the textual information.	熱い所のアスファルトに合ったらじゅわっと蒸発する感じか
	Evaluation	The reader makes comments or states an opinion about the text that is evaluative.	そりゃそうでしょう
	Reaction	The reader makes a comment to react, often emotionally to the text.	なつかしいな
Meta comment	Self-monitoring	The reader makes a comment about the degree of his/her own comprehension or use of a reading strategy.	うーん ちょっとよくわかんないな
	Text organization	The reader comments that refer to the structure of the text or the role of the information in the text.	一段落目は (中略) 主に一般的な例について挙げてて
Others	Others	The reader comments on other things that are irrelevant to his/her comprehension of the text.	ねむ…

Note. The verbal protocols were mostly commented in Japanese.