CrossMark

Children's integration of information across texts: reading processes and knowledge representations

Katinka Beker[1,2] · **Paul van den Broek[1]** · **Dietsje Jolles[1]**

Abstract

Constructing a knowledge representation from multiple texts requires the integration of information across texts. The aim of the current study was to investigate how elementary school students integrate information across multiple text passages and, particularly, whether students use information from a prior text to improve understanding of a current text. A sample of 105 children in grades 4 and 6 participated in the experiment. The multiple-text integration paradigm was used to study integration processes across texts during reading. Recall and (application) questions were used to investigate the extent to which information from different text passages was integrated into knowledge representations after reading. Individual differences in reading comprehension ability and working memory were also considered. The results indicate that children in both grades spontaneously activate information from an earlier text to aid their understanding during reading, and that they integrate information across texts in their knowledge representations. This was the case regardless of grade or individual differences in reading-comprehension ability and working memory. These findings provide insight into the mechanisms that may be involved in the integration of information across texts.

Keywords Intertextual integration · Multiple-text reading · Development · Reading processes · Knowledge representation

Introduction

Texts constitute one of the most important sources of information in education. Due to the growing quantity and availability of information on the Internet, learning and integrating information across *multiple* texts has become more and more common.

✉ Katinka Beker
k.beker@fsw.leidenuniv.nl

[1] Department of Educational Sciences, Leiden University, Leiden, The Netherlands

[2] Department of Education and Child Studies, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands

 Springer

This development poses challenges to learning that were previously restricted to expert readers (Goldman, 2015). These challenges need to be taken into account when designing school curricula, teacher instruction, and student assessment. Therefore, it is vital to improve our understanding of the skills and processes that are involved when readers construct a knowledge representation from multiple texts. The aim of the current study is to gain insight into how upper elementary school students integrate information across multiple text passages. Specifically, we investigate whether these children use information from a prior text to improve their understanding of a current text and, if so, whether such application of previously learned information is influenced by individual and developmental differences in reading-comprehension ability and working memory.

Integration processes and integration in memory

A common learning situation involves readers integrating information from several texts about one particular topic. For example, a reader may encounter an unfamiliar topic and then consult multiple webpages to learn more about that topic. More generally, when reading a new text on a topic about which he or she has already read in the past, the reader may recruit information from an earlier text to help understand the current text. In both cases, the information from the various texts needs to be connected and stored in a combined memory representation. Thus, processing of multiple texts crucially depends on the integration of information across the texts (e.g., Britt, Rouet, & Durik, 2018). The process of integration consists of two phases: (1) activating and integrating information across texts during reading, and (2) representing the integrated text information in memory (including relations across texts). Both aspects are important for achieving the educational standards that are relevant for learning from multiple texts in education (Common Core State Standards, 2010; OECD, 2015a, b).

During reading, each piece of information that is being processed activates associated information in memory, including information from previous parts of the same text and from background knowledge (Albrecht & O'Brien, 1993; McKoon & Ratcliff, 1992; van den Broek, Risden, Fletcher, & Thurlow, 1996). When reading multiple texts about the same topic, readers can also activate information from an earlier text when reading a later text (Beker, Jolles, Lorch, & van den Broek, 2016; Britt & Rouet, 2012; Perfetti, Rouet, & Britt, 1999), leading to co-activation of information from the two texts (Kendeou & O'Brien, 2014; Kintsch, 1988; van den Broek et al., 1996). As a result of this co-activation, a connection can be established between the elements of information from the two texts. Various types of connections are possible, ranging from purely associative to more meaningful connections, such as causal or logical relations. Whether a connection is made and, if so, whether it is integrated into a memory representation is determined by (1) the amount, and (2) the frequency of (co-)activation of information during reading (e.g., Kintsch, 1988; van den Broek et al., 1996). In the context of reading multiple texts, connections across texts may lead to an enriched knowledge representation or, in the case of conflicting information in the texts, to the incorporation into the knowledge

representation of information about sources and contradictions (Beker, Jolles, & van den Broek, 2017; Britt & Rouet, 2012).

It is clear—as every teacher will attest—that upper elementary school children *can* integrate information across different texts. But it is also clear that readers—both adults and children—often *fail* to recognize that there are connections between elements of information, even in single texts (Albrecht & O'Brien, 1993; Helder, Van Leijenhorst, & van den Broek, 2016; van der Schoot, Reijntjes, & van Lieshout, 2012). This is particularly the case when texts contain incomplete and (seemingly) inconsistent sections that can only be resolved by making an inference (i.e., by drawing conclusions based on the combined information across multiple texts). It has been well established that elementary readers have trouble reconciling inconsistencies within the same paragraph. It is likely that resolving inconsistencies by combining information across texts is even more difficult and may only occur when the young reader is prompted or guided by a teacher.[1] Given the correspondence between integration processes and the integration of information into memory, one would expect that a failure to integrate information into memory can be traced back to problems with integration processes during reading. Thus, it is important to study both the integration process and the resulting knowledge representation. In the next section we provide a brief review of previous investigations of integration processes and integrated knowledge representations in children.

Integration of text information

Three lines of research that focus on children's integration of information from texts are directly relevant for the present study. We summarize the main findings for each.

Integration across multiple auditorily presented texts

A first line of research concerns integration across texts by very young children, using auditorily presented texts (Bauer, King, Larkina, Varga, & White, 2012; Bauer & San Souci, 2010; Bauer, Varga, King, Nolen, & White, 2015). Bauer and colleagues had children aged 4–6 listen to story pairs that each included one stem fact (e.g., "groups of dolphins are called pods", "dolphins communicate by clicking and squeaking"). After a short time interval the children were asked questions that required them to integrate the stem facts after processing the materials (e.g., "how does a pod talk?"). The results showed that children as young as 4–6 years can integrate information from multiple (simple) texts, at least when prompted by the task and with auditorily presented texts. Although processing of auditorily presented

---

[1] Or when invited by the reading task. For example, after reading with a goal of knowledge gathering undergraduate readers were more likely to remember conflicting information across than within texts, presumably because the multiple document format prompted strategic processing aimed at cross-text coherence building (Stadtler, Scharrer, Brummernhenrich, & Bromme, 2013). It would be interesting to investigate whether elementary school children have these processing skills as well.

texts and that of written texts differ in many respects, these findings suggest that children in upper elementary school should be able to integrate information across different situations. Whether they actually will do so to improve understanding of an unclear or (seemingly) inconsistent text has not been established.

Integration within single texts

A second line of research focuses on integration of information within single texts. In line with the research presented above, several studies have indicated that children are able to integrate information within a single text (Cain & Oakhill, 1999; Oakhill, 1982, 1984), and that they make connections between text information and background knowledge (Barnes, Dennis, & Haefele-Kalvaitis, 1996; Cain, Oakhill, Barnes, & Bryant, 2001). It is important to note, however, that participants in these studies were explicitly prompted to combine and integrate information, for example by integration-promoting questions. The evidence shows that *spontaneous integration* is often more problematic. Although under some circumstances children are able to spontaneously integrate information during reading (Coté, Goldman, & Saul, 1998; Lynch & van den Broek, 2007; McMaster et al., 2012), it is often the case that children struggle with tasks that require spontaneous integration of information. For example, they have difficulty detecting internal inconsistencies (Markman, 1979; Oakhill, Hartt, & Samols, 2005), or repairing inconsistencies once detected (Helder et al., 2016; van der Schoot et al., 2012). These findings dovetail with those on adult readers that show that even experienced readers often fail to detect inconsistencies within a text unless they are prompted by specific reading goals or by textual cues that highlight the inconsistencies (e.g., Albrecht & Myers, 1995; Lea, Mulligan, & Walton, 2005). Integration across multiple texts is likely to be even be more challenging (Salmerón, Strømsø, Kammerer, Stadtler, & van den Broek, 2018), for example because the information is usually separated over a larger distance (Beker et al., 2016) or may be contradictory (Stadtler & Bromme, 2014); in addition, overlap between the texts may not be recognized (Kurby, Britt, & Magliano, 2005).

Integration across multiple printed texts

A third line of research focuses on the integration of information from multiple written texts. Studies involving adults and older adolescents suggest that comprehension and integration across multiple texts depends on a complex interplay of a number of different cognitive and motivational processes (for a recent review on on-line multiple-text integration processes, see Mason & Florit, 2018). Using a path analysis approach, for example, it has been demonstrated that multiple-text comprehension depends on prior knowledge, multiple-text comprehension strategies (i.e., strategies for comparing, contrasting, and integrating information across different texts), and effort; furthermore, that individual difference variables such as interest and *need for cognition* affect comprehension indirectly through their influence on strategy use and effort (Bråten, Anmarkrud, Brandmo, & Strømsø, 2014). Epistemic beliefs (i.e., a person's beliefs about the nature of knowledge and knowing) are also an important

factor influencing multiple-text comprehension (Bråten, Britt, Strømsø, & Rouet, 2011).

Little is known about the development of these cognitive and motivational factors in multiple text comprehension. One study has shown that young adolescents between the ages of 11 and 13 are at least *capable* of integrating information across texts during reading (Wolfe & Goldman, 2005). Using a think-aloud protocol during reading of conflicting historical texts, the authors showed that students engaged in self-explanation processes connecting information in the current text with prior knowledge, as well as with information that had just been read. Furthermore, there were individual differences in the degree to which students integrated information across texts, which was predictive of the level of subsequent reasoning about the texts. It is important to note that that the texts were specifically designed to promote inter-textual integration, and it is not clear whether students would engage in the same integration processes in more ambiguous situations. Moreover, it is possible that cross-text processing was encouraged by the use of the think-aloud procedure. It is well established that think-aloud procedures can be used to promote children's comprehension and learning (e.g., Walker, 2005). The important question of whether children *spontaneously* integrate multiple texts during reading is largely unexplored.

Although few studies have examined children's ability to integrate information across multiple texts *during* reading, their ability to relate information from different texts *after* reading has been studied relatively extensively and is included in international reading assessments such as PISA (OECD, 2018) and NAEP (Sheehan, Kostin, & Persky, 2006). A typical task involves students reading different texts and subsequently connecting information from the texts in response to a question. In general, these studies demonstrate that students struggle more with integrative questions than with questions that are non-integrative, for example questions that involve other reading skills such as locating information (e.g., Sabatini, O'Reilly, Halderman, & Bruce, 2014; Sheehan et al., 2006; Stadler & Bromme, 2014). Such findings indicate that reading different texts about the same topic does not necessarily lead to an integrated knowledge representation.

The current study

Integrating these three lines of research, it can be concluded that connecting information across texts is an important skill that children need to master in order to comprehend and learn from multiple texts, but that it is unclear whether children do so spontaneously. We experimentally investigated the spontaneous integration processes in young readers to shed light on factors that facilitate or hinder integration across multiple texts. By manipulating characteristics of texts, we can investigate fundamental integration processes across texts in a controlled setting. The aims of the current study were to determine: (1) whether upper elementary school children are able to resolve inconsistencies by integrating information across different text passages about the same topic, (2) whether they do so spontaneously, and (3) whether they incorporate intertextual connections (i.e., connections linking different

text passages) into memory. In order to answer these questions, the Multiple-Text Integration Paradigm (M-TIP) was used (Beker et al., 2016). In this paradigm participants read multiple short text passages about different topics. Texts about the same topic are presented in pairs, and the second text of a pair contains an *internal* inconsistency. There are two conditions. In the first condition the inconsistency can be resolved by activating an explanation from the first text. In the second condition the inconsistency cannot be resolved. Thus, the only difference between the conditions is whether the first text provides an explanation for the second text or not, so any difference in the processing of the second text can only be attributed to activation of information (i.e., the explanation) from the first text. Previous research with adults has demonstrated that the inconsistent target sentence in the second text is processed faster in the condition with explanations than in the condition without explanations (Beker et al., 2016). This speed-up indicates activation of information from the first text during reading of the second text, leading to co-activation of information from both texts. Several theoretical models state that co-activation of information leads to integration by forming a connection between the pieces of information that are co-active (Kendeou & O'Brien, 2014; Kintsch, 1988; van den Broek et al., 1996). The present study extends this work on adult readers by examining whether children also process the inconsistent target sentence faster in the condition with explanations than in the condition without explanations, which would indicate that they spontaneously activate information from a previous text while reading a subsequent text. This would be the case if activation of prior information occurs relatively automatically, as a result of *spread of activation*, triggered by overlapping concepts in the two texts (Cook & O'Brien, 2014; Kintsch, 1988). Alternatively, if effortful processing is required to process inconsistencies, it may be difficult for elementary school children to resolve the inconsistencies, especially for those students whose working memory capacity and reading comprehension ability are limited. As this study is among the first to investigate children's implicit, spontaneous integration processes across different texts, we purposely kept the distance between consecutive text passages small, omitted source information, and kept the inconsistencies obvious. By providing optimal conditions for integrating information across texts, we establish a baseline that allows for comparisons with situations in which integrating information across different texts becomes more challenging (e.g., greater distance between conflicting information or more subtle inconsistencies).

In order to examine whether children incorporate intertextual connections into their knowledge representation, we asked children to recall the texts. Recall can be useful to gain insight into knowledge representations built from multiple texts (Britt & Sommer, 2004). Children were asked to report everything they remembered from the text, without interference by the experimenter. We used a general measure of the knowledge representation because we were interested in spontaneous integration of information across texts and we therefore did not want to prompt deliberate integration across texts. By identifying the source of each information unit, we determined to what extent information from multiple texts was integrated into the knowledge representation: The number of switches *between* the texts was taken as a measure of integration. If children demonstrate activation of prior text information during reading of a subsequent text (as indicated by a difference in reading times between the

condition with and without explanation), one would expect this to be reflected in the knowledge representation, because co-activation of prior and current text information may lead to constructing or strengthening a connection between the two co-activated elements (van den Broek et al., 1996). In analyzing the recalls, the focus was on indications of integration, but because more integration may also have a positive effect on overall memory for the texts, a measure of total recall was also included. This allows us to test the possibility that the hypothesized effects of conditions are a byproduct of higher recall.

Individual and developmental differences

Single-text processing studies have demonstrated that integrative processes are more difficult for children with poor reading comprehension skills and for children with low working-memory abilities (Cain & Oakhill, 2007; Helder et al., 2016; Long, Oppy, & Seely, 1997; McMaster et al., 2012; van der Schoot et al., 2012). Measures of reading comprehension ability and working memory were included in the current study to determine whether they interact with integration across multiple texts during reading. If reading-comprehension ability and the availability of working-memory capacity contribute to integration across texts, then one would expect a stronger effect of condition (with versus without explanations) for children with good reading-comprehension skills and good working-memory skills than for children with poor reading comprehension skills and poor working memory skills; condition would be expected to affect both the reading time measures and measures of intertextual connections in memory.

In the current study we included children in grade 4 of elementary school because they have mastered the basic reading skills and because—as prescribed by the national educational standards (Expertisecentrum Nederlands [Expertise Centre Netherlands], 2010)—at this age children are expected to integrate information across texts at least at a basic level. Many skills related to comprehension monitoring and detecting/repairing inconsistencies continue to develop from childhood into adulthood (Kendeou, van den Broek, White, & Lynch, 2009; Oakhill & Cain, 2012). Therefore, we also included children from grade 6 to determine whether there are grade-related differences in the ability to resolve inconsistencies using information from prior texts. Based on previous studies we expected main effects of grade on reading times (Fuchs & Fuchs, 1993) and integration (Bauer & San Souci, 2010). Finally, we expected that the older children would have stronger integration skills and, therefore, that the effects of conditions with and without explanations would be stronger for children in grade 6 than for children in grade 4.

Method

Participants

The research sample consisted of 105 children from Grade 4 ($N=54$ with 30 girls and 24 boys, Mean age$=9.9$, $SD=0.4$) and Grade 6 ($N=51$ with 30 girls and 21

boys, Mean age = 11.9, *SD* = 0.4) from four regular Dutch primary schools. The study was approved by the university's ethical review board and informed consent was obtained from the parents. Only children with good or corrected eyesight and without developmental and reading disorders were included in the experiment to make sure that they were able to manage the task demands. Participation was rewarded with a small gift.

Materials and design

Text materials

We created a child-friendly version of the multiple-text integration paradigm (M-TIP; Beker et al., 2016). Children read expository text[2] pairs, in which the second text contained an internal inconsistency, and the first text either contained or did not contain an explanation that could help resolve the inconsistency in the second text[3] (the Inconsistent-with-explanation and Inconsistent-without-explanation conditions, respectively). The conditions were manipulated within subjects. The texts used in prior research were adapted to fit the reading level of children in Grades 4 and 6. To check whether the difficulty level of the adapted texts was appropriate for children in Grades 4 and 6, a readability index was used that provides an indication of the difficulty of the texts based on a variety of text characteristics [the (Dutch) Cito readability index for primary education, or P-CLIB; Evers, 2008; Staphorsius, Verhelst, & Kleintjes, 1996]. The average readability index score of the adapted texts indicated that the texts were appropriate for children in Grades 4 and 6.

A pilot study was conducted with a different group of 4th and 6th grade children to determine whether the inconsistencies in the resulting set of texts were salient for children this age and whether the explanations were convincing. The reason for this pilot study was that previous research has shown that internal inconsistencies sometimes are not detected by children this age (August, Flavell, & Clift, 1984; Markman, 1979). Based on the results of this pilot study, we only included texts with inconsistencies that were salient enough to be detected by the majority of the children and explanations that were plausible enough to be linked to the inconsistency.

The topics of the expository texts were realistic but fictitious, to limit the influence of background knowledge. There were 20 different topics (i.e., text pairs, so 40 texts per child) concerning animals, persons, objects, countries, and events. The topics were based on real-world situations (e.g., the text about the 'rulver' was based on the polar fox). For each topic there were two versions of each text pair, which were counterbalanced across subjects: A text with an inconsistency in combination with a preceding text that contained an explanation for the inconsistency, and a text with an inconsistency in combination with a preceding text that did *not* contain an explanation for the inconsistency. In the condition with explanation, the first text

---

[2] Expository texts were used because it is common to use this genre to present new ideas (Singer, 2015).

[3] Note that the two texts within each pair always were consistent with each other.

described an explanation that could resolve the inconsistency. In the condition without explanation, the first text described additional information about the topic that could not resolve the inconsistency. The texts with inconsistencies were the same in both conditions. The texts had an average length of 8 sentences. The inconsistency was manifested in the target sentence, which was always the penultimate sentence of the text. The target sentences were between 50 and 53 characters in length. Example materials are presented in Table 1.

The texts within a pair were designed to be independent: each formed a syntactically and semantically complete text and could be comprehended in isolation (with the exception of the part with the inconsistency in the second text). Every text began with an introductory sentence and ended with a closing sentence, and each concept was introduced as if it were new. This was expected to increase the awareness among readers that they are reading multiple texts and not just paragraphs of a single text.

Questions

The children received three types of questions. The first type of question (comprehension question) was a multiple-choice question with two alternatives (yes/no). The purpose of this type of question was twofold: (1) to test whether children were paying attention to the task and (2) to indicate when the child had finished reading the text. The question always concerned literal information from the preceding text and was the same in all conditions. The second type of question (recall question) was an open question about the main topic of the text. The question always followed the same format: "What do you remember from the text about topic X?", where X represents the main topic of the text pair (often the fictitious animal/object/person, for example the 'rulver'). The third type of question (application question) also was an open question. The purpose of this question was to create a task that stimulates reading for learning. These questions always introduced a problem in a novel setting that required the application of the explanation from the text. For example, in the rulver text the application question was: "Imagine walking in a natural history museum. You are walking past all sorts of mounted animals. Suddenly you see two rulvers, one brown rulver and one white rulver. Why do you think they have a different color?"

Working memory

Children completed a translated version of the sentence-span task of working memory (originally created by Daneman & Carpenter, 1980; adapted by Swanson, Cochran, & Ewers, 1989). This task involved the processing and storage of sentences and words. Children listened to sets of unrelated sentences, answered a comprehension question about one of these sentences, and then recalled the last word of each sentence. There were six levels that increased in difficulty, with each level consisting of two sets of items. The items at the easiest level consisted of two sentences whereas the items at the most difficult level consisted of six sentences. There were 10 sets in total. The task was stopped either when a child was not able to answer the comprehension question correctly or when she/he was not able to recall

Table 1 Example text materials showing two versions of the topic 'the rulver'

	Inconsistent-with-explanation	Inconsistent-without-explanation
Text 1	The rulver is an animal with a short tail	The rulver is an animal with a short tail
	The rulver lives mainly on the moors, but sometimes also in the woods	The rulver lives mainly on the moors, but sometimes also in the woods
	The rulver has a pretty brown fur	The rulver has a pretty brown fur
	This fur can be used to make clothing	This fur can be used to make clothing
	Hunters can get a lot of money for this fur	Hunters can get a lot of money for this fur
	In the winter the rulver's fur turns white	*That is why they try to catch rulvers*
	Its brown fur fell off in the fall	*They catch fewer rulvers than they used to*
	After this, new white hairs start to grow	*Because there are not many rulvers left*
	White camouflage is better against the snow	*The hunters are not happy about this*
Text 2	The rulver's fur can be used to make coats	The rulver's fur can be used to make coats
	To get this fur, the rulver is being hunted in the summer	To get this fur, the rulver is being hunted in the summer
	The rulver's fur has a special brown color	The rulver's fur has a special brown color
	You don't see this brown color on other animals	You don't see this brown color on other animals
	In the winter the hunt for the rulver stops	In the winter the hunt for the rulver stops
	Because then you cannot see the rulver in the white snow	Because then you cannot see the rulver in the white snow
	The hunt can resume in June	The hunt can resume in June

The differences between first texts in the Inconsistent-with-explanation and Inconsistent-without-explanation condition are italicized. The underlined word is what makes the underlined target sentence inconsistent (in the Inconsistent-with-explanation and Inconsistent-without-explanation conditions). These sample texts are translated from Dutch

at least one word in each set within one level. The final score was calculated as the total number of questions answered correctly and the total number of words recalled correctly (regardless of the order in which the answers were given). Fourth graders scored on average 6.02 points (SE = .58) and sixth graders scored on average 7.58 points (SE = .57). The performance difference between groups was not significant ($t(103) = -1.94$, $p = .056$).

Reading-comprehension ability

The Cito test for reading comprehension is a national standardized, norm-referenced test (Cito, 2013). In this test, children read a variety of texts and have to answer multiple-choice questions about these texts. Cito reading comprehension tests are administered twice each year in each grade to assess children's reading comprehension skills. Performance scores of the Cito test for reading comprehension for Grades 4 and 6 were obtained from the teachers of the children. The most recent test results were used. On average, the test was administered 2 months before the experiment. Test results are reported as 'level' scores, which consist of five levels ranging from I (i.e., the highest level) to V (the lowest level). Each level represents 20% of the range of norm scores. These levels indicate the level of reading-comprehension ability based on norms from a large sample of children of the same age. The majority (90–95%) of the schools in the Netherlands used the Cito test for reading comprehension at the time of testing, so the norms are representative (Egberink, Janssen, & Vermeulen, 2015). The Cito assessment for reading comprehension in Grades 4 and 6 has good reliability and validity (Egberink et al., 2015). The majority of the children in fourth grade (80%) and sixth grade (90%) scored at or above the national average on the reading comprehension test. This indicates that good comprehenders were overrepresented in our sample. The average score of fourth graders was 36.11 (SE = 2.26) and the average score of sixth graders was 66.71 (SE = 1.64). This difference was significant ($t(103) = -11.03$, $p < .001$).

Procedure

Children first received verbal instructions about the procedure of the reading task. They were told that they were going to read texts sentence-by-sentence. They were asked to read the texts for comprehension and to answer several questions about them. The comprehension questions were asked immediately after reading the texts, the recall and application questions after a delay (i.e., after reading four text pairs). Half of the children received a hint about the relatedness of the text pairs. Because the presence/absence of a hint did not influence any of the measures of interest, this factor was left out of the analyses.

After the verbal instructions, children were asked to read the same instructions on the screen, and they performed two practice trials. If necessary the experimenter gave feedback during the practice trials. When children demonstrated comprehension of the task during the practice trials, they were instructed to continue to the remainder of the experiment individually and feedback was no longer provided.

Before each text was presented, the message "next text" was presented in the center of the display screen to indicate the beginning of a new text and thereby increasing the boundary between texts that were part of a pair and between texts with different topics. This message was presented in capital letters to increase the awareness that children were going to read a new text that was distinct from the previous text. The next screen showed a fixation cross in the center of the screen that was presented for a variable interval of between 500 and 2500 ms before each sentence. Following this fixation cross, sentences were presented one by one in the center of the screen. Children were instructed to read at their own pace. They could progress to the next sentence by pressing the space bar. To prohibit children from skipping a sentence by accidentally double-hitting the space bar, the program did not respond to a press if it occurred within 500 ms of the previous press. Also, if children took longer than 15,000 ms to read a sentence the program automatically continued to the next sentence. After reading each text, children were presented with one comprehension question. The children were instructed to keep their thumbs on the space bar, and their index fingers on the "yes" and "no" keys at all times (the "S" and "L" keys on the keyboard). They did not receive feedback about the accuracy of their answers. The order in which the text pairs were presented was counterbalanced across subjects, and the order in which the texts that belonged to one pair was presented was fixed, with the text with the inconsistency always immediately following the text with or without explanation (but as with each text, separated by a question and the message "next text"). After reading four text pairs, the children were asked to answer the recall questions. The recall questions were presented in the same order as the topics were presented to the children in the texts. Children were instructed to report only the most important information from the text. In case of a nonresponse (no response or "I don't know") the experimenter asked a question (e.g., "don't you remember anything about topic X?") to elicit a response. After each recall question, the application question was asked. In case of a nonresponse (e.g., silence or "I don't know") the experimenter told the child that they were allowed to use their imagination. When children only said yes or no, the experimenter asked why. Children were asked to report their answers verbally and their responses were recorded with an audio tape recorder.

Each testing session consisted of two parts, each part lasting about 35 min on average, with a break in between. Children were told that their participation was voluntary and that at any time they could quit-none did. Instructors were instructed to be alert for signs of fatigue. All children reported that they enjoyed participating in the experiment; this was in agreement with the impressions by their teachers and the instructors. Ten children had additional breaks during the experiment due to (unexpected) obligations at school. Additional breaks always took place after a block of four texts pairs and the corresponding questions, to make sure that the time delay between reading and answering questions was similar for all blocks in all children.

Coding and scoring

Recall

Children's auditory responses were transcribed, parsed into idea units, and each idea unit was coded. An idea unit generally comprised a semantically meaningful clause (consisting of a subject and main verb). Each unit was coded based on the most likely origin of the information: (1) the first text of the pair, (2) the second text of the pair, (3) both texts, (4) background knowledge. Non-meaningful, incomplete clauses (e.g., "he was…[silence]") and metacognitive responses (e.g., "I don't remember") were excluded from the analysis. Next, the number of source switches between the first and the second text was counted, for responses that could be traced to one unique text (i.e., code types 1 and 2). For example, a recall response consisting of 4 idea units coded as coming from texts 1-1-2-2 would result in an integration score of 1, because there is one switch between sources (i.e., the second idea unit, which relates to the first text, is followed by an idea unit from the second text of a pair). To determine reliability of scoring, 25% of the responses were coded by two raters (the first author and several trained faculty members). The remaining responses were coded by the first author only. Agreement between the raters was good (Mean Cohen's κ = 0.68).

Application questions

Responses to the application questions were coded as 'correct' when children used (parts of) the explanation from the first text in a pair to answer the question, and 'incorrect' when they gave a different response. Two raters (the first author and a trained faculty member) coded 25% of the responses to the application questions. The remaining answers were coded by the first author only. Agreement between the raters was good (Cohen's κ = 0.69).

Results

Reading times

Before analyzing the data, the responses to the questions and the reading times were inspected. On average, children answered 87% of the questions correctly, which demonstrates that the children were processing the texts. Reading times that deviated over 2.5 standard deviations on both the subject and item means were removed, because these were assumed to reflect processes that are not of interest in the current study (Ratcliff, 1993). Less than 1% of the data was removed using this criterion. Descriptive statistics are presented in Table 2.

As the distribution of the reading times was skewed to the right, the reading times were transformed by taking the natural log of each score to make the distribution more symmetrical (Richter, 2006). Because of the multilevel structure of the data (Richter, 2006), reading times were analyzed using hierarchical linear models using

Table 2 Descriptive statistics for each condition in each grade (reading times in milliseconds)

	Condition	M	SE
Grade 4[a]	Inconsistent-with-explanation	3654.34	73.90
	Inconsistent-without-explanation	3791.60	74.01
Grade 6[b]	Inconsistent-with-explanation	3360.23	60.51
	Inconsistent-without-explanation	3607.30	66.87
Difference between grades	$t(103) = .72, p = .476$		
Correlation with working memory	$r = -.14, p = .156$		
Correlation with reading ability	$r = -.23, p = .016$		

[a]$N = 54$

[b]$N = 51$

R-statistics software and the 'lmerTest', 'effects' and 'MuMIn' packages. Item-level reading speeds were represented at the first level and subjects and items (texts) were represented at the second level of the model, with the items nested within conditions. Subjects and items were treated as random effects, whereas the predictors (Condition, Grade, Reading Comprehension Ability, and Working Memory) were treated as fixed factors. Continuous predictors (i.e., Working Memory) were centered on the grand mean. Degrees of freedom were estimated with Satterthwaite's approximation method (Kuznetsova, Brockhoff, & Christensen, 2015; SAS Technical Report R-101, 1978; Satterthwaite, 1941). Effects were classified as significant when $p < .05$. Restricted maximum likelihood was used to fit the models. The model was built in two steps. In the first step a model that included Condition was compared to a model without predictors (i.e., the baseline model) by statistically testing the improvement in model fit using likelihood ratio tests. The addition of the factor Condition significantly improved the model compared to the baseline model ($\chi^2(1) = 15.73, p < .001$). The mean reading time of the target sentence in the Inconsistent-with-explanation condition was significantly faster than the mean reading time of the target sentence in the Inconsistent-without-explanation condition ($b = .05$). In the second step, the main effects of the background variables (Grade, Reading Comprehension Ability, and Working Memory) and the two-way interactions between Condition and each background variable were added to the model that only included Condition to determine whether the effect of Condition was qualified by an interaction with the background variables. The background variables and interactions did not significantly improve the model ($\chi^2(6) = 8.11, p = .230$). An overview of the model comparisons is presented in Table 3.[4]

---

[4] We also conducted the analyses excluding trials with incorrect responses to the comprehension questions. The results were the same: The predictor Condition was a significant addition to the model ($\chi^2(1) = 11.71, p < .001$). The addition of the background variables did not explain variance over and above the effect of Condition ($\chi^2(6) = 6.44, p = .38$).

Table 3 Model comparisons

Model	Tested against	Model fit			
		Reading times	Integration	Total recall	Application[a]
Baseline	Condition	$\chi^2(1)=15.73^*$ $R^{2b}_{GLMM(m)}=.004$ $R^{2c}_{GLMM(c)}=.509$	$\chi^2(1)=16.98^*$ $R^{2b}_{GLMM(m)}=.008$ $R^{2c}_{GLMM(c)}=.250$	$\chi^2(1)=.01$ $R^{2b}_{GLMM(m)}<.001$ $R^{2c}_{GLMM(c)}=.032$	–
Condition	Condition + Condition * RCA + Condition * WM + Condition * Grade	$\chi^2(6)=8.11$ $R^{2b}_{GLMM(m)}=.019$ $R^{2c}_{GLMM(c)}=.516$	$\chi^2(6)=22.63^*$ $R^{2b}_{GLMM(m)}=.045$ $R^{2c}_{GLMM(c)}=.253$	$\chi^2(6)=16.95^*$ $R^{2b}_{GLMM(m)}=.042$ $R^{2c}_{GLMM(c)}=.326$	$\chi^2(3)=43.43^*$

RCA reading comprehension ability; *WM* working memory. All models contain a random intercept over persons and items. The model fit measures reflect comparisons between the two models in the left two columns. The asterisk indicates an interaction between predictors

[a]For the application measure the variable Condition was excluded from the model because only the responses in the Inconsistent-with-explanation condition were taken into account

[b]This represents the proportion of variance explained by the fixed factor(s) alone. It is a measure of the effect size (Johnson, 2014; Nakagawa & Schielzeth, 2013)

[c]This represents the proportion of variance explained by both the fixed and the random factors. It is a measure of the effect size (Johnson, 2014; Nakagawa & Schielzeth, 2013)

*$p < .01$

Recall

In the majority of cases (77%) the answers in response to recall questions contained information from the text pairs. When children were not able to recall information from the text pairs, this was probably due to insufficient cues to recollect the information: Each recall question contained only one non-specific recall cue (e.g., 'the animal') in combination with the unfamiliar topic (e.g., 'the rulver'). This cue may not always have been sufficient for the child to recall which of the four preceding unfamiliar topics had to be retrieved. There is one indication that supports this possibility: When children did not report content-specific information in response to the recall question, they *did* recall text information spontaneously in response to the subsequent application question for an additional 12% of the questions, possibly because these application questions contained additional cues. Because the application questions did not explicitly prompt recall and, therefore, not all children took the opportunity to report what they remembered after listening to the application question, the recall analyses were based on the responses to the recall questions only.

The descriptive statistics are presented in Table 4. The integration scores were analyzed using hierarchical linear models following the same procedures and steps as in the previous analyses (Table 3). The addition of the factor Condition contributed significantly to the model compared to the baseline model ($\chi^2(1)=16.98$, $p<.001$). The integration score was higher in the Inconsistent-with-explanation condition than in the Inconsistent-without-explanation condition ($b=.19$). Addition of the background variables and interactions significantly improved the model ($\chi^2(6)=22.63$, $p<.001$). In particular, Reading Comprehension Ability was positively related to integration scores ($t(169)=3.94$, $b=.167$). However, this effect was

Table 4 Descriptive statistics for each condition in each grade (recall data)

Condition	Integration[a]		Total Recall[b]	
	M	*SE*	*M*	*SE*
Grade 4[c]				
Inconsistent-with-explanation	1.01	.06	4.13	.14
Inconsistent-without-explanation	.92	.05	4.21	.15
Grade 6[d]				
Inconsistent-with-explanation	1.23	.05	4.86	.15
Inconsistent-without-explanation	.96	.05	4.73	.15
Difference between grades	$t(103) = -1.51, p = .134$		$t(103) = -1.69, p = .094$	
Correlation with working memory	$r = .13, p = .183$		$r = .75, p < .001$	
Correlation with reading ability	$r = .39, p < .001$		$r = .35, p < .001$	

[a]The score represents the mean integration scores on each topic

[b]The score represents the mean number of recalled idea units on each topic

[c]Number of data points = 401 (inconsistent-with-explanation) and 403 (inconsistent-without-explanation)

[d]Number of data points = 405 (inconsistent-with-explanation) and 405 (inconsistent-without-explanation)

not significant after correction for total recall ($p = .330$). There were no other significant main or interaction effects.

Total recall was analyzed using hierarchical linear models using the same procedures as in the previous analyses (Table 3). Condition did not contribute significantly to the model compared to the baseline model ($\chi^2(1) = 0.01$, $p = .909$). However, the background variables and interactions significantly improved the model ($\chi^2(6) = 16.95$, $p = .009$). In particular, Reading Comprehension Ability was positively related to total recall ($t(130) = 3.40$, $b = .53$). There were no other main or interaction effects.

Application questions

The primary purpose of the application questions was to create a task that stimulates reading for learning. However, the responses to these questions are also of interest, particularly to explore the potential effects of individual differences in the background variables. Application scores were analyzed using logistic hierarchical linear models, using the same model-building procedures as in the previous analyses (Table 3). Only the responses in the Inconsistent-with-explanation condition were analyzed, because only these questions could be answered by applying the knowledge from both texts in a pair. The background variables explained a significant amount of variance of application scores ($\chi^2(3) = 43.43$, $p < .001$). In particular, there was a main effect of Reading Comprehension Ability: The ability to comprehend texts was positively related to application scores ($z = 5.60$, $b = .46$, $p < .001$). There was also a main effect of Grade ($z = 3.08$, $b = .46$ $p = .018$): Children in sixth grade performed better on the application questions ($M_{proportion\_correct} = .55$, $SE = .02$)

than children in fourth grade ($M_{proportion\_correct} = .42$, $SE = .02$). Working memory did not affect the performance on application questions.

Discussion

An important goal in education is to learn from multiple texts (Britt et al., 2018; Common Core State Standards, 2010; Salmerón et al., 2018). This requires the processing of individual texts, as well as the integration and encoding of information from multiple texts. If learning is successful, the knowledge representation constructed from multiple texts can be used to solve new problems. In the current study two aspects of learning from multiple texts were investigated in primary school children: the learning process and the resulting knowledge representation. The research questions were (1) whether upper elementary school children (grades 4 and 6) are able to resolve inconsistencies by integrating information across multiple text passages; (2) whether they do so spontaneously; and (3) whether they incorporate intertextual connections (i.e., connections linking different texts) into their memory. In investigating these questions, we also considered possible effects of differences in reading comprehension ability, working memory, and grade.

Integration across texts during reading

The multiple-text integration paradigm (M-TIP) was used to determine whether information from previous texts was spontaneously activated and used to resolve inconsistencies during reading of subsequent texts (Beker et al., 2016). As hypothesized, the processing speed of inconsistent target sentences in subsequent texts was faster when prior texts contained explanations for the inconsistencies than when prior texts lacked explanations. Thus, in the condition with explanations, information from the current and the previous text was available at the same time during reading. This co-activation of current and previous text information may enable the reader to create connections across texts (Kendeou & O'Brien, 2014; Kintsch, 1988; van den Broek et al., 1996). These results show that children as young as 9 attempt to relate information across texts by spontaneously activating information from previous texts during the reading of subsequent texts. This is in line with what has been observed in adults, using the same paradigm (Beker et al., 2016), and in older children (aged 11–13), using think-aloud methods (Wolfe & Goldman, 2005). The present results extend previous findings by showing, using an unobtrusive measure, that integration across texts occurs *spontaneously* during reading (Bauer et al., 2012, 2015; Bauer & San Souci, 2010; Wolfe & Goldman, 2005).

Although the current results seem to conflict with previous studies that showed that children particularly struggled with integrating information across texts (Sabatini et al., 2014; Sheehan et al., 2006), there are important differences between the current study and previous studies that may explain the seemingly contradictory conclusions. First, whereas previous studies used explicit questions that required the production of responses, we used an implicit measure to inspect spontaneous

integration of information across texts. Second, in the current study we created optimal conditions for the integration of information across texts (i.e., by using clear inconsistencies). In previous studies the conditions may have been more challenging (i.e., by using longer texts with a mix of obvious and non-obvious inconsistencies). Successfully integrating information across texts is likely to depend on situational circumstances. Future studies should focus on manipulating different aspects of the situation to determine under what circumstances integrating information across texts becomes more challenging. By systematically increasing the difficulty of the materials, for example by increasing the temporal distance or by decreasing the conceptual overlap between the elements of information to be connected, a clear picture could emerge of when and why children sometimes fail to integrate information across texts.

The current findings raise the question whether co-activating information actually led the young readers to integrate the information in a meaningful way. It is possible that overlap in key terms between the first and the second text might lead to activation of information from the first text, but that this might lead only to an associative connection and not a meaningful connection (such as a causal relation, e.g., 'the rulver is difficult to see in the white snow *because* it changes color in the winter'). Future research could employ think-aloud methods in combination with the multiple-text paradigm to determine whether co-activated information is indeed connected by the reader and, if so, whether the relation is meaningful (for example, causal), associative, or both.

Constructing a knowledge representation from multiple texts

The knowledge representation of the texts was analyzed by asking children to recall as much as they could from the texts. As hypothesized, children showed more integrated recall in those situations where connecting the two texts could restore comprehension, i.e., in the conditions that provided explanations as opposed to the conditions that lacked explanations. Processing times of the target sentence suggest that integration during recall was the result of co-activation of information during reading. This is in line with current theoretical models of the integration process during reading (Cook & O'Brien, 2014; Kendeou & O'Brien, 2014; Kintsch, 1988; van den Broek & Kendeou, 2008; van den Broek et al., 1996). Importantly, the effect was not a byproduct of higher recall in general, because on total recall there were no differences between the conditions. Interestingly, in prior research adult readers did not show a condition difference in the integration of information into their knowledge representation (Beker et al., 2016). In that study, a different, possibly less sensitive coding procedure was used, which raises the possibility that if the current procedure were used, adults too might show conditions differences. Therefore, it would be useful to investigate the apparent discrepancy between adults and children by including different measures of knowledge representations (e.g., primed-recognition measures, which directly probe connections in the memory representation, see Myer & O'Brien, 1998; van den Broek & Lorch, 1993), and to directly compare adults with children on these measures using the same materials. Recall procedures such

as the one employed in the current study have some limitations (e.g., selectivity in what a participant reports) that may be obviated by using (a combination of) other measures.

Individual differences in integration across texts

Our results show that children were able to resolve the inconsistencies using information from a prior text. In contrast to our predictions, the ability to resolve inconsistencies was not affected by grade, working memory, or reading ability. Therefore, it is possible that the development of the ability to resolve inconsistencies using prior information is less protracted than that of other aspects of processing internal inconsistencies. One possible explanation is that the activation of background information about the same topic (including the solution to the inconsistency) occurs relatively automatically, through a process of spread of activation (Cook & O'Brien, 2014; Kintsch, 1988). Therefore, the inconsistency could be resolved quickly, and the reader might not even notice that there was an inconsistency. If background knowledge about the apparent inconsistency is not available, however, the reader may have to engage in more effortful processing, leading to longer reading times. It may be that this aspect of processing internal inconsistencies shows a more protracted development. It is important to note, however, that the lack of correlations with reading skill and working memory may also reflect ceiling effects. The texts were intentionally constructed to be as comprehensible as possible for purposes of inviting integration of information across texts. Moreover, the majority of the participants in this study scored well on the comprehension test. Thus, variability in comprehension skills was relatively small. For these reasons the strongest and weakest comprehenders in our study may have demonstrated similar behaviors. Furthermore, it may be that the relatively small distance between the texts enabled both low- and high-span readers to keep information from the first text activated. Finally, the lack of a grade effect may also be due to the simplicity of the task. The current study was intentionally designed to minimize the challenges posed by the separate texts, in order to encourage learning from multiple texts (Beker et al., 2016). This may be why differences across grades were negligible. It is possible that more challenging multiple text situations allow for a wider range of (strategic) processes, which may differentiate children in different developmental stages. Future research should address this possibility; this would increase our knowledge about the boundary conditions that determine success or failure in multiple text situations.

Individual differences in transfer

As hypothesized, reading comprehension ability and grade affected readers' ability to apply information from a text to a new situation (i.e., transfer). Good comprehenders performed better on this task than did poor comprehenders. There are several explanations for this finding. Good comprehenders may have constructed better knowledge representations of the texts than poor comprehenders (Oakhill, 1982), or their knowledge representations may have been more available, which helped them

in answering the application questions. Furthermore, children in Grade 6 generally performed better than did children in Grade 4, suggesting that the ability to transfer develops over time. This is consistent with other research on the development of transfer skills (Thibaut & French, 2016).

Mechanisms involved in integration processes

The difference in reading time observed between the experimental conditions may reflect a speed-up in the condition with explanation, or it may reflect a reduced slow-down. Without a baseline measure we cannot distinguish between these accounts, but we can speculate about the direction of the effect on the basis of previous research. There are (at least) two possibilities: The effect can be explained in terms of inconsistency resolution or in terms of pre-activation. According to the inconsistency resolution account, the inconsistency in the target sentence is detected, and this then triggers activation of previous text and background information. In the condition with explanation this would lead to a reduced slow-down, because activation of the explanation from the first text helps resolve the inconsistency. In the condition without explanation, the inconsistency is believed to trigger an (unsuccessful) memory search, resulting in longer processing times. According to the pre-activation account, the information from previous parts of the text and background knowledge is already activated before the reader processes the target sentence, for example due to featural overlap (Myers & O'Brien, 1998; van den Broek et al., 1996). In the condition with explanation this is posited to lead to increased efficiency in processing the target sentence because it readily fits prior knowledge. In this case, the reader may not even experience an inconsistency. In the condition without explanation this is believed to lead to longer processing times because the target sentence does not fit the knowledge representation. Recent insights in the field of predictive inferences favor the pre-activation account (for a review, see Kutas, DeLong, & Smith, 2011). Furthermore, a previous study using the multiple-text integration paradigm demonstrated that the processing speed of the inconsistent target sentence in the condition with explanation was comparable to the processing speed of the same target sentence in a consistent situation, providing further support for the pre-activation account (Beker et al., 2016). Whatever the mechanism that leads to activation of prior text information, both accounts explain how information from prior texts is activated during reading the target sentence, enabling co-activation of information from both texts, and possibly integration. The accounts differ only in when co-activation begins: before or during the reading of the target sentence. In terms of theoretical models of reading comprehension, it would be important to gain more insight into the fundamental processes that underlie integration across multiple texts (Britt et al., 2018; Salmerón et al., 2018).

Directions for future research

In the current study, several cues were used to increase the distinctive boundary between the two texts: an intervening task (a comprehension question); an explicit

message ("next text"); implicit text structure cues (e.g., introducing the topic such as the 'rulver' in the second text as if it were new); and, for half the children, hints that each text was part of a pair (e.g., "You are going to read *two texts* in a row. When reading the *second text*, try to think of the *first text*"). Despite these cues, it is possible that children did not always perceive the texts in one pair as distinct. The M-TIP paradigm can easily be extended to study spontaneous integration processes during reading in situations in which integration is more challenging for children, for instance in more naturalistic text settings, such as when reading texts from the Internet. This is especially relevant given that the processes that we target in this study are of crucial importance in and outside school. Thus, the current study provides a foundation for future investigations of integration of information across texts in children in a controlled, experimental manner in more ecologically valid situations. As mentioned, a fruitful direction for future research would be to vary the difficulty of activation and integration across texts in order to determine which factors affect intertextual integration. Candidate factors include those that have been found to affect integration within single texts (e.g., featural overlap, reading strategies, intra-text distance, etc.), but also factors that are particularly relevant in the context of multiple texts (e.g., textual/physical distance between texts, reliability of the sources, differing writing styles, etc.).

In the current study, recall was used as a measure for knowledge representation. Sometimes children in our study were not able to recall information from the text, possibly due to insufficient cues to remember a topic that is unfamiliar to them. Even when they did recall information from the text, the total number of idea units recalled was rather low: between 4 and 5 idea units per text. Given the brevity of the texts (8 sentences), the instruction to recall only the most important information, and the unfamiliarity of the topics to the children, such low recall is not surprising. As a result, the number of connections that can be made across texts is also limited, making it difficult to establish the effects of the conditions and individual differences on integrating information across texts. By exploring the issues addressed in this study during the reading of longer, more complex texts, we can establish boundary conditions for successful cross-text integration and potential interactions with individual differences.

Conclusion

It has been argued that learning from multiple texts may be difficult for children, for example when children do not recognize the relatedness of the texts (Bauer et al., 2012; Kurby et al., 2005), when the distance between the texts is large (Beker et al., 2016), or when children are taught to process texts in isolation from other texts (Hartman & Hartman, 1994). However, the results of the current study suggest that upper elementary school children are *capable* of processing texts in relation to other texts, and that they can do so spontaneously, without explicit prompting. Children demonstrated integrative processing across texts during reading, and integrated information from different texts into their memory representation. It is important to note that, by using short text passages that were read directly after one another, we

created optimal circumstances for integration across texts. In doing so, our results provide a first step towards gaining more insight into the process of learning from multiple texts and can be used as a starting point to reveal factors that facilitate or inhibit learning from multiple texts.

References

Albrecht, J. E., & Myers, J. L. (1995). The role of context in accessing distant information during reading. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 21,* 1459–1468.

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

August, D. L., Flavell, J. H., & Clift, R. (1984). Comparison of comprehension monitoring of skilled and less skilled readers. *Reading Research Quarterly, 20*(1), 39–53.

Barnes, M. A., Dennis, M., & Haefele-Kalvaitis, J. (1996). The effects of knowledge availability and knowledge accessibility on coherence and elaborative inferencing in children from six to fifteen years of age. *Journal of Experimental Child Psychology, 61*(3), 216–241.

Bauer, P. J., King, J. E., Larkina, M., Varga, N. L., & White, E. A. (2012). Characters and clues: Factors affecting children's extension of knowledge through integration of separate episodes. *Journal of Experimental Child Psychology, 111*(4), 681–694.

Bauer, P. J., & San Souci, P. (2010). Going beyond the facts: Young children extend knowledge by integrating episodes. *Journal of Experimental Child Psychology, 107*(4), 452–465.

Bauer, P. J., Varga, N. L., King, J. E., Nolen, A. M., & White, E. A. (2015). Semantic elaboration through integration: Hints both facilitate and inform the process. *Journal of Cognition and Development, 16*(2), 351–369.

Beker, K., Jolles, D., Lorch, R. F., Jr., & van den Broek, P. (2016). Learning from texts: Activation of information from previous texts during reading. *Reading and Writing, 29,* 1161–1178.

Beker, K., Jolles, D., & van den Broek, P. (2017). Meaningful learning from texts: The construction of knowledge representations. In J. A. Léon & I. Escudero (Eds.), *Reading comprehension in educational settings* (pp. 29–62). Amsterdam: John Benjamins.

Bråten, I., Anmarkrud, Ø., Brandmo, C., & Strømsø, H. I. (2014). Developing and testing a model of direct and indirect relationships between individual differences, processing, and multiple-text comprehension. *Learning and Instruction, 30,* 9–24.

Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J. F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist, 46*(1), 48–70.

Britt, M. A., & Rouet, J.-F. (2012). Learning with multiple documents: Component skills and their acquisition. In M. J. Lawson & J. R. Kirby (Eds.), *Enhancing the quality of learning: Dispositions, instruction, and learning processes* (pp. 276–314). New York: Routledge.

Britt, M. A., Rouet, J. F., & Durik, A. (2018). *Literacy beyond text comprehension: A theory of purposeful reading*. New York: Routledge.

Britt, M. A., & Sommer, J. (2004). Facilitating textual integration with macro-structure focusing tasks. *Reading Psychology, 25*(4), 313–339.

Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing, 11*(5–6), 489–503.

Cain, K., & Oakhill, J. (2007). Reading comprehension difficulties: Correlates, causes, and consequences. In K. Cain & J. Oakhill (Eds.), *Children's comprehension problems in oral and written language: A cognitive perspective* (pp. 41–76). New York, NY: Guilford.

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition, 29*(6), 850–859.

Cito. (2013). *LOVS Begrijpend lezen groep 8 [Reading comprehension Grade 6]*. Arnhem: Cito.

Common Core State Standards. (2010). *Common core state standards for english language arts & literacy in history/social studies, science, and technical subjects*. Retrieved from http://www.corestandards.org.

Cook, A. E., & O'Brien, E. J. (2014). Knowledge activation, integration, and validation during narrative text comprehension. *Discourse Processes, 51*(1–2), 26–49.

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

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

Egberink, I. J. L., Janssen, N. A. M., & Vermeulen, C. S. M. (2015). *COTAN beoordeling 2015, Leerling- en onderwijsvolgsysteem begrijpend lezen [COTAN review 2015, student and education monitoring and evaluation system reading comprehension]*. Retrieved from www.cotandocumentatie.nl.

Evers, G. (2008). *Programma voor berekening Cito LeesIndex voor het Basisonderwijs [Program for calculating Cito Reading Index for primary education] P-CLIB versie 3.0*. Arnhem: Cito.

Expertisecentrum Nederlands [Expertise Centre Netherlands]. (2010). *Doorlopende Leerlijnen [Continuing learning standards]*.

Fuchs, L. S., & Fuchs, D. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22*(1), 27.

Goldman, S. R. (2015). Reading and the web: Broadening the need for complex comprehension. In R. J. Spiro, M. DeSchrijver, M. S. Hagerman, P. M. Morsink, & P. Thomas (Eds.), *Reading at a crossroads? Disjunctures and continuities in current conceptions and practices* (p. 89). New York: Routledge.

Hartman, J. A., & Hartman, D. K. (1994). *Arranging multi-text reading experiences that expand the reader's role (Tech. Report No. 604)*. Retrieved from Urbanan-Champaign: University of Illinois: https://www.ideals.illinois.edu/bitstream/handle/2142/17586/ctrstreadtechrepv01994i00604_opt.pdf.

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.

Johnson, P. C. D. (2014). Extension Nakagawa & Schietzeth's R_GLMM^2 to random slopes models. *Methods in Ecology and Evolution, 5,* 44–946.

Kendeou, P., & O'Brien, E. J. (2014). The Knowledge Revision Components (KReC) framework: Processes and mechanisms. In D. N. Rapp & J. L. G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences*. Cambridge, MA: MIT Press.

Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology, 101*(4), 765–778.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163–182.

Kurby, C. A., Britt, M. A., & Magliano, J. P. (2005). The role of top-down and bottom-up processes in between-text integration. *Reading Psychology, 26*(4–5), 335–362.

Kutas, M., DeLong, K. A., & Smith, N. J. (2011). A look around what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future* (pp. 190–207). New York: Oxford University Press.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). *Package 'lmerTest'*. Retrieved from https://cran.r-project.org/web/packages/lmerTest/lmerTest.pdf.

Lea, R. B., Mulligan, E. J., & Walton, J. L. (2005). Accessing distant premise information: How memory feeds reasoning. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*(3), 387–395.

Long, D. L., Oppy, B. J., & Seely, M. R. (1997). Individual differences in readers' sentence- and text-level representations. *Journal of Memory and Language, 36*(1), 129–145.

Lynch, J. S., & van den Broek, P. (2007). Understanding the glue of narrative structure: Children's on- and off-line inferences about characters' goals. *Cognitive Development, 22*(3), 323–340.

Markman, E. M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child Development, 50*(3), 643–655.

Mason, L., & Florit, E. (2018). Complementary methods for assessing online processing of multiple sources. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of multiple source use* (pp. 425–446). New York: Routledge.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99*(3), 440.

McMaster, K. L., van den Broek, P., Espin, C. A., White, M. J., Rapp, D. N., Kendeou, P., et al. (2012). Making the right connections: Differential effects of reading intervention for subgroups of comprehenders. *Learning and Individual Differences, 22*(1), 100–111.

Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes, 26*(2–3), 131–157.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from Generalized Linear Mixed-Effects Models. *Methods in Ecology and Evolution, 4,* 133–142.

Oakhill, J. (1982). Constructive processes in skilled and less skilled comprehenders' memory for sentences. *British Journal of Psychology, 73,* 13–20.

Oakhill, J. (1984). Inferential and memory skills in childrens comprehension of stories. *British Journal of Educational Psychology, 54,* 31–39.

Oakhill, J., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading, 16*(2), 91–121.

Oakhill, J., Hartt, J., & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and Writing, 18,* 657–686.

OECD (2015a). *OECD skills outlook 2015: Youth, skills and employability*. Retrieved from: http://www.oecd.org/education/oecd-skills-outlook-2015-9789264234178-en.htm.

OECD (2015b). *PISA 2015 results (volume I): Excellence and equity in education*. Retrieved from http://www.oecd.org/education/pisa-2015-results-volume-i-9789264266490-en.htm.

OECD (2018). *PISA 2018 draft analytical frameworks*. Retrieved from https://www.oecd.org/pisa/data/PISA-2018-draft-frameworks.pdf.

Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Toward a theory of documents representation. In H. Van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 99–122). Mahwah, NJ: Lawrence Erlbaum Associates.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114*(3), 510–532.

Richter, T. (2006). What is wrong with ANOVA and multiple regression? Analyzing sentence reading times with hierarchical linear models. *Discourse Processes, 41*(3), 221–250.

Sabatini, J. P., O'Reilly, T., Halderman, L., & Bruce, K. (2014). Broadening the scope of reading comprehension using scenario-based assessments: Preliminary findings and challenges. *L'Année Psychologique, 114*(4), 693–723.

Salmerón, L., Strømsø, H. I., Kammerer, Y., Stadtler, M., & van den Broek, P. (2018). Comprehension processes in digital reading. In P. van den Broek (Ed.), *Learning to read in a digital world* (pp. 91–120). Amsterdam: John Benjamins.

SAS Technical Report R-101. (1978). *Tests of hypotheses in fixed-effects linear models*. Retrieved from http://support.sas.com/documentation/onlinedoc/v82/techreport_r101.pdf.

Satterthwaite, F. E. (1941). Synthesis of variance. *Psychometrika, 6*(5), 309–316.

Sheehan, K. M., Kostin, I., & Persky, H. (2006). Predicting item difficulty as a function of inferential processing requirements: An examination of the reading skills underlying performance on the NAEP Grade 8 reading assessment. In *Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA.*

Singer, M. (2015). Validation of text and discourse inferences- and explicit content. In E. J. O'Brien, A. E. Cook, & R. F. Lorch Jr. (Eds.), *Inferences during reading* (pp. 68–93). Cambridge: Cambridge University Press.

Stadtler, M., & Bromme, R. (2014). The content-source integration model: A taxonomic description of how readers comprehend conflicting scientific information. In D. N. Rapp & J. L. G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 379–402). Cambridge, MA: MIT Press.

Stadtler, M., Scharrer, L., Brummernhenrich, B., & Bromme, R. (2013). Dealing with uncertainty: Readers' memory for and use of conflicting information from science texts as function of presentation format and source expertise. *Cognition and Instruction, 31,* 130–150.

Staphorsius, G., Verhelst, N. D., & Kleintjes, F. G. M. (1996). De ontwikkeling van een domeingerichte index voor leesbaarheid en leesvaardigheid. [The development of an index for readability and reading skills]. *Taalbeheersing, 18,* 116–132.

Swanson, H. L., Cochran, K. F., & Ewers, C. A. (1989). Working memory in skilled and less skilled readers. *Journal of Abnormal Child Psychology, 17,* 145–156.

Thibaut, J.-P., & French, R. M. (2016). Analogical reasoning, control and executive functions: A developmental investigation with eye-tracking. *Cognitive Development, 38,* 10–26.

van den Broek, P., & Kendeou, P. (2008). Cognitive processes in comprehension of science texts: The role of co-activation in confronting misconceptions. *Applied Cognitive Psychology, 22,* 335–351.

van den Broek, P., & Lorch, R. F., Jr. (1993). Network representations of causal relations in memory for narrative texts: Evidence from primed recognition. *Discourse Processes, 16,* 75–98.

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). Hillsdale, NJ: Erlbaum.

van der Schoot, M., Reijntjes, A., & van Lieshout, E. C. (2012). How do children deal with inconsistencies in text? An eye fixation and self-paced reading study in good and poor reading comprehenders. *Reading and Writing, 25*(7), 1665–1690.

Walker, B. J. (2005). Thinking aloud: Struggling readers often require more than a model. *The Reading Teacher, 58,* 688–692.

Wolfe, M. B. W., & Goldman, S. R. (2005). Relations between adolescents' text processing and reasoning. *Cognition and Instruction, 23*(4), 467–502.